

# 中华人民共和国国家标准

GB/T 28043—2011/ISO 13528:2005

## 利用实验室间比对进行能力验证的统计方法

Statistical methods for use in proficiency testing by interlaboratory comparisons

(ISO 13528:2005, IDT)

2011-10-31 发布

2012-02-01 实施



中华人民共和国国家质量监督检验检疫总局  
中国国家标准化管理委员会

发布

中 华 人 民 共 和 国  
国 家 标 准  
利用实验室间比对进行能力验证的统计方法  
GB/T 28043—2011/ISO 13528:2005

\*

中国标准出版社出版发行  
北京市朝阳区和平里西街甲2号(100013)  
北京市西城区三里河北街16号(100045)  
网址 [www.spc.net.cn](http://www.spc.net.cn)  
总编室:(010)64275323 发行中心:(010)51780235  
读者服务部:(010)68523946  
中国标准出版社秦皇岛印刷厂印刷  
各地新华书店经销

\*

开本 880×1230 1/16 印张 3.75 字数 110 千字  
2012年3月第一版 2012年3月第一次印刷

\*

书号: 155066·1-44297

如有印装差错 由本社发行中心调换  
版权专有 侵权必究  
举报电话:(010)68510107

## 目 次

|                                                         |     |
|---------------------------------------------------------|-----|
| 前言 .....                                                | III |
| 引言 .....                                                | IV  |
| 1 范围 .....                                              | 1   |
| 2 规范性引用文件 .....                                         | 1   |
| 3 术语和定义 .....                                           | 1   |
| 4 设计和解释实验室能力验证的统计指南(见 GB/T 15483.1—1999,5.4.2) .....    | 2   |
| 4.1 行动信号和警戒信号 .....                                     | 2   |
| 4.2 对指定值不确定度的限定 .....                                   | 2   |
| 4.3 重复测量次数的确定 .....                                     | 3   |
| 4.4 样本的均匀性与稳定性(见 GB/T 15483.1—1999,5.6.2 和 5.6.3) ..... | 3   |
| 4.5 定义测量方法 .....                                        | 3   |
| 4.6 数据报告(见 GB/T 15483.1—1999,6.2.3) .....               | 4   |
| 4.7 能力验证结果的有效期 .....                                    | 4   |
| 5 确定指定值及其标准不确定度 .....                                   | 4   |
| 5.1 确定指定值的方法选择 .....                                    | 4   |
| 5.2 配方法(见 GB/T 15483.1—1999,A.1.1,a) .....              | 4   |
| 5.3 有证参照值(见 GB/T 15483.1—1999,A.1.1,b) .....            | 5   |
| 5.4 参照值(见 GB/T 15483.1—1999,A.1.1,c) .....              | 5   |
| 5.5 专家实验室的公议值(见 GB/T 15483.1—1999,A.1.1,d) .....        | 7   |
| 5.6 参加者的公议值(见 GB/T 15483.1—1999,A.1.1,e) .....          | 7   |
| 5.7 指定值的比对 .....                                        | 11  |
| 5.8 缺失值 .....                                           | 12  |
| 6 确定能力评定标准差(见 GB/T 15483.1—1999,A.2.1.3) .....          | 12  |
| 6.1 方法选择 .....                                          | 12  |
| 6.2 由规定值确定 .....                                        | 12  |
| 6.3 由经验预期值确定 .....                                      | 12  |
| 6.4 由一般模型确定 .....                                       | 13  |
| 6.5 由精密度试验结果确定 .....                                    | 14  |
| 6.6 由一轮能力验证计划所得数据确定 .....                               | 14  |
| 6.7 能力验证得到的精密度与测量方法已知精密度的比较 .....                       | 14  |
| 7 性能统计量的计算 .....                                        | 15  |
| 7.1 实验室偏倚的估计(见 GB/T 15483.1—1999,A.2.1.4,a) .....       | 15  |
| 7.2 百分相对差(见 GB/T 15483.1—1999,A.2.1.4,b) .....          | 17  |
| 7.3 秩与秩百分数(见 GB/T 15483.1—1999,A.2.1.4,c) .....         | 19  |
| 7.4 $z$ 值(见 GB/T 15483.1—1999,A.2.1.4,d) .....          | 21  |
| 7.5 $E_n$ 值(见 GB/T 15483.1—1999,A.2.1.4,e) .....        | 23  |

|              |                                                   |    |
|--------------|---------------------------------------------------|----|
| 7.6          | $z'$ 值                                            | 23 |
| 7.7          | $\zeta$ 值                                         | 24 |
| 7.8          | $E_z$ 值                                           | 24 |
| 7.9          | 已报告不确定度时进行数据分析的示例                                 | 24 |
| 7.10         | 组合性能统计量值                                          | 28 |
| 8            | 一轮能力验证计划中多个被测量的组合性能统计量值的图示法                       | 28 |
| 8.1          | 应用                                                | 28 |
| 8.2          | 性能统计量值的直方图                                        | 28 |
| 8.3          | 标准化实验室偏倚的条形图                                      | 29 |
| 8.4          | 标准化重复性测量值的条形图                                     | 30 |
| 8.5          | 尧敦图                                               | 30 |
| 8.6          | 重复性标准差图                                           | 36 |
| 8.7          | 分割样品(见 GB/T 15483.1—1999, A.3.1.2)                | 38 |
| 9            | 组合多轮能力验证计划中性能统计量值的图示法(见 GB/T 15483.1—1999, A.3.2) | 41 |
| 9.1          | 应用                                                | 41 |
| 9.2          | $z$ 值的常规控制图                                       | 41 |
| 9.3          | $z$ 值的累积和控制图                                      | 43 |
| 9.4          | 实验室标准化偏倚对平均值的图                                    | 43 |
| 9.5          | 点图                                                | 44 |
| 附录 A (规范性附录) | 符号                                                | 46 |
| 附录 B (规范性附录) | 样品的均匀性检验和稳定性检验                                    | 47 |
| 附录 C (规范性附录) | 稳健分析                                              | 50 |
| 参考文献         |                                                   | 52 |



## 前 言

本标准按照 GB/T 1.1—2009 给出的规则起草(技术内容除外)。

本标准等同采用国际标准 ISO 13528:2005《利用实验室间比对进行能力验证的统计方法》。对 ISO 13528:2005 的错误作了如下修改和更正:

- 将 4.3 第三段中“重复性标准差”和“能力评定标准差”分别更正为“重复性方差”和“能力评定方差”;
- 将 5.6.1 第二段中“式(53)”更正为“式(C.1)”;
- 将 7.1.3 中“图 1 和 2”更正为“图 2 和图 3”;
- 将 7.2.3 中“图 1 和 3”更正为“图 2 和图 4”;
- 将 7.4.3 中“图 1 和 4”更正为“图 2 和图 5”;
- 将 7.7.2 中“ $X \pm U_x$ ”更正为“ $\bar{x} \pm U_x$ ”;
- 将 7.8 第二段中的“7.4 或 7.6”更正为“5.4 或 5.6”,将“7.5”更正为“5.5”;
- 将 7.9.2 中“ $u_x = 1.23 \times s^* / \sqrt{181} = 13 \times 10^{-10}$ ”更正为“ $u_x = 1.25 \times s^* / \sqrt{181} = 13 \times 10^{-10}$ ”;
- 将 7.9.4 中“ $p > 16$ ”更正为“ $p > 17$ ”;
- 将附录 A 中“ $s_x$ ”更正为“ $s_{\bar{x}}$ ”;
- 将附录 B.2 中第四行和第五行的“标准差”更正为“方差”;
- 将附录 B.3 中式(B.4)和式(B.7)中“ $x_{i..}$ ”均更正为“ $\bar{x}_{i..}$ ”。

本标准的附录 A、附录 B 和附录 C 均为规范性附录。

本标准由全国统计方法应用标准化技术委员会(SAC/TC 21)提出并归口。

本标准起草单位:中国标准化研究院、中国合格评定国家认可中心、中国科学院数学与系统科学研究院、北京工业大学、山东检验检疫局技术中心、辽宁出入境检验检疫局。

本标准主要起草人:张帆、丁文兴、田玲、谢田法、于振凡、翟培军、冯士雍、郭武、郑江、陈志民。

## 引 言

### 0.1 能力验证的目的

利用实验室间比对的能力验证用于确定实验室对特定测试或测量的实施能力,并监测实验室的持续执行能力。能力验证目的的详细阐述可见 GB/T 15483.1 的引言。从统计的角度来说,实验室能力往往用三个特性来刻画:实验室偏倚、稳定性和重复性。实验室偏倚和重复性在 GB/T 3358.1、GB/T 3358.2和 GB/T 6379.1 中已定义。实验室测试结果的稳定性可由 GB/T 6379.3 定义的中间精密度来度量。

如果有标准物质,可用 GB/T 6379.4 中的程序对标准物质测试来评估实验室偏倚。另外,利用实验室间比对的能力验证提供了获得实验室偏倚信息的一般方法,用能力验证的数据来估计实验室偏倚也是对此类数据进行分析的一个重要方面。因为稳定性和重复性也会影响能力验证获得的数据,所以,实验室可利用一轮能力验证计划的实验数据,来辨别偏倚实际上是由较差的稳定性还是较差的重复性造成的。因此对实验室能力的这几个方面进行定期评估是非常重要的。

稳定性评估通过对保留样本的再测试获得,或通过对标准物质或内部标准物质(实验室自身储备的物质,作为专用标准物质)的定期测量获得,该方法已在 GB/T 6379.3 中给出。在能力验证中,也可通过绘制控制图来评估实验室的稳定性,这是分析能力验证数据的另一个重要方面。但这种方法在单轮能力验证计划中是不可行的。

适用于重复性评估的数据可由实验室日常工作中获得,或由专门用于评估重复性而进行的实验得到。因此,重复性评估不是能力验证的一个重要方面。尽管检测实验室重复性还是非常重要的。可利用 GB/T 6379.6 中给出的极差控制图方法来评估重复性。

图 1 给出了本标准中统计技术应用的流程图。

### 0.2 GB/T 15483

GB/T 15483.1 描述了不同类型的能力验证计划,并给出了组织和设计能力验证计划的指南。GB/T 15483.2 给出了实验室认可机构选择和使用能力验证计划的指南。这两项标准应作为能力验证领域的参考文件(内容并未与本标准重复)。GB/T 15483.1 附录中简要描述了能力验证计划中统计方法。

本标准是 GB/T 15483.1 的补充,给出了能力验证中统计方法使用的详细指南。本标准在很大程度上是基于分析实验室能力验证的协调草案<sup>[1]</sup>而设计,但对于所有的测量方法都适用。

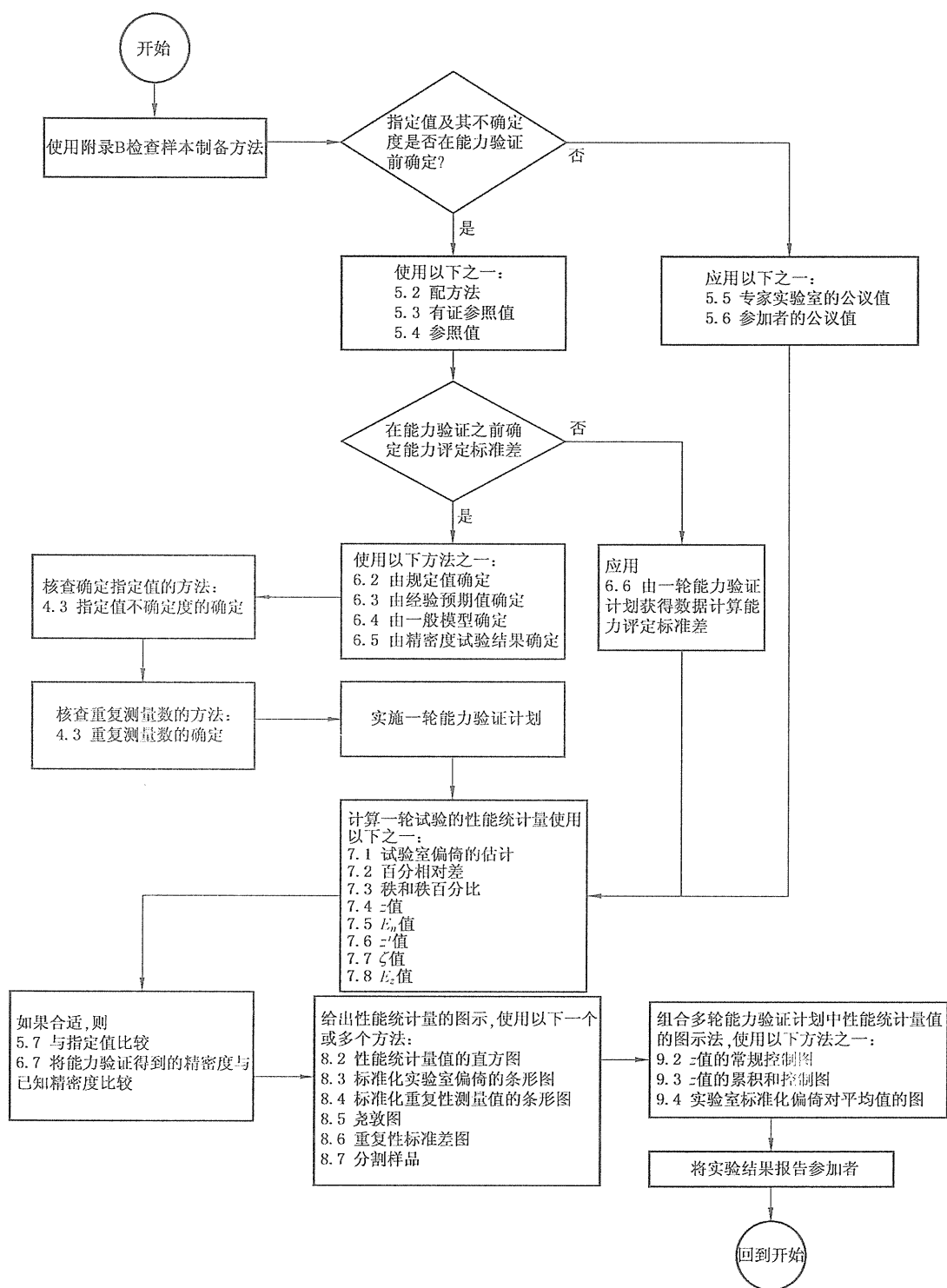


图1 实施能力验证计划时应用统计技术的流程图

# 利用实验室间比对进行能力验证的统计方法

## 1 范围

本标准规定了在能力验证计划中数据分析的统计方法,并且对能力验证计划参加者和认可机构在实际工作中使用上述方法提出了建议。

本标准适用于通过实验室的测量结果验证其不存在明显的不可接受的偏倚。

本标准适用于定量数据而不适用于定性数据。

## 2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 3358.1—2010 统计学词汇及符号 第1部分:一般统计术语与用于概率的术语 (ISO 3534-1:2006, IDT)

GB/T 3358.2—2010 统计学词汇及符号 第2部分:应用统计 (ISO 3534-2:2006, IDT)

GB/T 6379.1—2004 测量方法与结果的准确度(正确度与精密度) 第1部分:总则与定义 (ISO 5725-1:1994, IDT)

GB/T 15483.1—1999 利用实验室间比对的能力验证 第1部分:能力验证计划的建立和运作 (ISO/IEC Guide 43-1:1997, IDT)

## 3 术语和定义

GB/T 3358.1、GB/T 3358.2、GB/T 6379.1界定的以及下列术语和定义适用于本文件。

### 3.1

**实验室间比对** **interlaboratory comparison**

在预定条件下,对两个或两个以上实验室就同一或相似的检测对象进行检测或测量的组织、实施和评价。

### 3.2

**能力验证** **proficiency testing**

利用实验室间比对确定实验室检测能力。

### 3.3

**指定值** **assigned value**

对于给定目的具有适当不确定度的赋予特定量的值,有时该值是约定采用的。

### 3.4

**能力评定标准差** **standard deviation for proficiency assessment**

基于可用信息,用于能力评估的离散性度量。

### 3.5

**z 值** **z-score**

由能力验证的指定值和标准差计算的实验室偏倚的标准化度量。

注:有时也称为  $z$  比分数。

## 3.6

协调者 coordinator

能力验证计划中负责所有协调工作的组织机构或个人。

## 4 设计和解释实验室能力验证的统计指南(见 GB/T 15483.1—1999, 5.4.2)

## 4.1 行动信号和警戒信号

4.1.1 本标准给出了一些简单的数值或图形准则,它将应用于能力验证所得数据,以判断这些数据能否引发行动信号或警戒信号。即使在一个运作良好、有经验丰富工作人员的实验室,偶尔也会得到异常数据;同样地,即使一个已由精密度试验验证为有效的标准测量方法也有可能存在缺陷,而这个缺陷可能只在多轮能力验证后才显现;并且能力验证计划本身也可能存在某些缺陷。由于以上原因,此处给出的准则不应作为处罚实验室的依据。若能力验证被用于处罚实验室,则有必要以此为目的设计出适合的准则。

4.1.2 该准则设计的目的在于,当能力评定标准差是基于观测结果(使用 6.4 至 6.6 中的方法之一)时,该准则可在观测结果极其异常,需要进一步调查并采取纠正措施的情形下发出行动信号。

4.1.3 协调者应了解能力验证数据中可预见变异的主要来源。任何分析的第一步应对测量结果的分布进行检查,以发现未预见的变异来源。例如,若出现双峰分布则表明测量结果可能来自于混合总体,其原因可能是测量方法不同、样本被污染或测量方法操作指南不清晰造成的。这种情况下,应先解决有关问题再进行分析 and 评估。实验室认可机构应事先制定应对能力验证中出现不可接受结果的政策,后续措施由此政策或实验室质量管理办法来确定。然而,当能力验证中实验室产生了不可接受结果时仍有一般性推荐步骤,4.1.4 中给出了具体的行动指南。

4.1.4 在能力验证中,当能力评定标准差由测量结果得出时,若有结果给出行动信号,实验室应采取适当的调查和纠正措施,如有必要,可与协调者或实验室认可机构协商。除非有正当理由,实验室应检查其自身程序,并向实验员确认纠正措施以避免此类结果再次出现。实验室可向协调者询问该问题产生的可能原因,或请协调者咨询其他专家。实验室应参加之后轮次的能力验证计划,以评估其纠正措施的有效性。适当的纠正措施可能是以下几项之一:

- a) 核查相关人员是否理解并遵循测量程序;
- b) 核查测量程序的所有细节是否正确;
- c) 核查设备校准和试剂的成分;
- d) 更换可疑的设备或试剂;
- e) 与另一个实验室进行人员、设备和/或试剂的比对测试。

GB/T 15483.2—1999 中描述了实验室认可机构对能力验证结果的使用。

## 4.2 对指定值不确定度的限定

指定值  $X$  的标准不确定度  $u_x$ ,依赖于  $X$  的定值方法,并且,当该值由若干实验室的测试结果得出时,则  $u_x$  依赖于实验室数或可能的其他因素。第 5 章给出了指定值标准不确定度的计算方法。

能力评定标准差  $\hat{\sigma}$  用来评估能力验证中实验室偏倚的大小。计算能力评定标准差的方法在第 6 章给出,它与实验室偏倚的估计值进行比较的准则在第 7 章给出。

当指定值的标准不确定度  $u_x$  远大于能力验证中使用的能力评定标准差  $\hat{\sigma}$  时,会存在一种风险,即某些实验室将会因为指定值不准确而收到行动或警戒信号,而不是因为实验室内部的任何原因。基于以上理由,应确定指定值的标准不确定度,并报告给参加能力验证计划的实验室(见 GB/T 15483.1—1999, A.1.4 和 A.1.6)。

当

$$u_x \leq 0.3\hat{\sigma} \quad \dots\dots\dots (1)$$

时,指定值的不确定度可忽略,并可不包含在能力验证结果中。

当以上准则不满足时,协调者应考虑以下几点:

- a) 寻找一种确定指定值的方法,使指定值的不确定度满足以上准则。
- b) 在能力验证的结果解释中使用指定值的不确定度(见 7.5 关于  $E_n$  值或 7.6 关于  $z'$  值有关描述)。
- c) 通知能力验证的参加者,指定值的不确定度不可忽略。

示例:假定指定值  $X$  由 11 个实验室测试结果的平均值  $\bar{x}$  确定,能力评定标准差为这 11 个结果的标准差  $s$ ,则  $\hat{\sigma}=s$ 。

作为估计,指定值的标准不确定度可由  $u_x=s/\sqrt{11}\approx 0.3s$  确定,显然符合以上准则。然而,当实验室数少于 11 个时则不满足该准则。进一步地,当样本是非均匀或不稳定的,或者存在导致实验室结果共同偏倚的因素(例如,当他们使用同样的参照标准)时,指定值的不确定度将大于  $s/\sqrt{11}$ 。

### 4.3 重复测量次数的确定

重复性变异是造成能力验证中实验室偏倚间差异的原因之一。当重复性变异比能力验证的标准差大得多时,重复性变异有可能导致出现无规律的能力验证结果。这种情况下,实验室可能在某一轮试验中存在较大偏倚,但下一轮中不会存在较大偏倚,且很难确定原因。

因此,当决定有必要限制重复性变异的影响时,能力验证中每个实验室的重复测试次数  $n$  应满足:

$$\sigma_r/\sqrt{n} \leq 0.3\hat{\sigma} \quad \dots\dots\dots (2)$$

其中  $\sigma_r$  是在此之前由实验室间实验确定的重复性标准差。

使用系数 0.3 是因为,当以上条件满足时,重复性标准差对能力评定标准差的贡献(按方差计)不多于 10%。

所有实验室应进行相同次数的重复测量。(本标准给出的分析方法中假定这个条件已满足。)若不等式(2)的条件不满足,则应增加重复测量次数,或者对能力验证结果给出提醒。

该方法假定实验室具有大体相似的重复性。当此条件不满足时将产生其他情况。这些情况下,当应用本标准所述方法时,可能会用到以下技巧。协调者可利用重复性标准差的典型值来确定重复测量次数  $n$ 。进而,各实验室应检查它本身的重复性标准差是否满足不等式(2)。若不满足,则应调整其测量程序(通过增加测试次数,测试结果的平均为一个测量结果),使结果满足不等式(2)。

### 4.4 样本的均匀性与稳定性(见 GB/T 15483.1—1999,5.6.2 和 5.6.3)

附录 B 给出了用于检验能力验证的样本是否具备充分的均匀性和稳定性的方法。

当使用的某种样本制备方法不能满足附录 B 中的均匀性检验准则时,参加者应测试重复样本,或者确定能力评定标准差应考虑样本不均匀性,如附录 B 所述。

### 4.5 定义测量方法

对于定义测量方法,测量结果与测量方法密切相关。例如,颗粒物的粒度分布可能由方孔或圆孔筛决定。可能无法确定哪种筛分法更优,除非指定某种筛分法,实验室应用不同的筛分法可能会得到不同的结论。如果参加者使用了一种不同于确定指定值的方法时,即使测量过程无任何过错时他们的结果也可能发生偏倚。

如果参加者可自由选择测量方法,则他们不会产生任何有效的一致结果。有两种办法可以克服这个问题:

- a) 当参加者按常规使用标准方法时,应用此方法建立指定值,并且指导其他参加者在能力验证中也使用这个方法。

b) 对所使用的每一种方法各建立一个单独的指定值。

当指定了被测量但未指定测量程序时,将产生类似的情况,也必须做出相同的选择。

#### 4.6 数据报告(见 GB/T 15483.1—1999,6.2.3)

由于能力验证中数据计算的需要,建议单个测量结果的舍入误差不大于  $\sigma_r/2$ 。

应要求参加者报告其测量结果的实际值。测量结果不应表示为区间值(例如,不应将结果报告为“ $<0.1$ ”或“小于检出限”)。同样的,当观测值为负时,则须报告实际的负数值(例如,即使逻辑上测量结果不能为负数时也不应将结果报告为0)。应告知参加者,若其报告了某样本的区间值,或将负数结果报告为0时,则在分析时应剔除此样本的所有数据。如有必要,报告结果所用表格可包含一栏允许参加者指明结果是否低于检出限。

#### 4.7 能力验证结果的有效期

实验室在单独一轮能力验证计划中所得结果的有效期限定在实验室本轮操作的时间内。因此,如果实验室在单独一轮计划中获得了一个满意的结果,并不表明其他情况下该实验室也可获得可靠的数据。

那些按质量体系运作,并且曾经在多轮次能力验证计划中获得了满意结果的实验室,应认定该实验室能够获得持续可靠的数据。

### 5 确定指定值及其标准不确定度

#### 5.1 确定指定值的方法选择

在5.2至5.6中描述了五种确定指定值  $X$  的方法。按照 GB/T 15483.1 的要求,协调者应负责方法的选择,并遵循技术专家的咨询结果。当参加计划的实验室数很少时,5.5和5.6中所述方法可能不适用。本章给出的计算指定值标准不确定度  $u_x$  的方法通常足以满足在本标准中的应用。当然,其他方法也可使用,只要其有可靠的统计学基础并且能力验证计划文件中对此方法有详细描述。

协调者应负责指定值的确定。指定值不应透露给参加者,直到他们已将自己的实验结果报告给了协调者。协调者应出具一份报告包含以下细节:如何获得指定值,参与确定指定值的实验室,对指定值的溯源性和测量不确定度的说明。

JJF 1059—1999《测量不确定度评定与表示》对测量不确定度的评定给出了指导。

当稳健统计方法是最合适方法时,本标准推荐使用该方法(如5.5和5.6中所示)。或者,也可使用包含了检验和剔除离群值的方法,只要它有可靠的统计学基础且方法已文件化。GB/T 6379.2给出了检验离群值的指南。

#### 5.2 配方法(见 GB/T 15483.1—1999,A.1.1,a)

##### 5.2.1 总则

被测物可能由特定比例的混合成分组成,或是添加一定比例的成分到基质中得到。这种情况下,指定值  $X$  由各组分的量值计算所得。

按照配方法直接制备试样是非常有用的。此时只要保证材料成分的比例或是添加成分的比例,不需准备集样,同时也保证了检测材料的均匀性。然而,当由配方法得到的样本中添加物比典型材料要更松散,或者呈现不同的形态,则应使用其他更好的方式获得样本。

##### 5.2.2 指定值的标准不确定度 $u_x$

当指定值由测试物质的配方来计算时,根据 JJF 1059—1999《测量不确定度评定与表示》,用不确定度分量来合成标准不确定度。例如,在化学分析中,不确定度通常与重量和体积有关。

本方法(在化学分析中)的局限性在于需要保证以下几点:

- a) 基质不受添加成分的影响,或者添加成分在基质中的比例是精确已知的;
- b) 所有成分混合均匀(当有此要求时);
- c) 所有误差来源是已被识别(例如,人们往往意识不到玻璃可以吸收汞化合物,实际上汞化合物水溶液的浓度会因容器材质而改变);
- d) 添加成分和基质之间没有反应。

### 5.2.3 示例:已硬化混凝土中水泥成分的确定

本例中,对混凝土成分的测试,可先对各成分(水泥,沙石和水)称重,再混合并形成混凝土样本。由于已知所制备样本的准确成分,该方法比用分析法确定水泥成分要好得多。

## 5.3 有证参照值(见 GB/T 15483.1—1999, A.1.1, b)

### 5.3.1 总则

若能力验证中所用物质为有证标准物质(CRM),则该物质的有证参照值就可作为指定值  $X$ 。

注1:有证参照值有时也称为有证参考值。

注2:有证标准物质有时也有证标准样品。

### 5.3.2 指定值的标准不确定度 $u_x$

当有证标准物质用作被测材料时,指定值的标准不确定度由证书中给出。

该方法的局限性在于,为能力验证中的每个参加者都提供一份有证标准物质会比较昂贵。

### 5.3.3 示例:骨料的 LA 值

“LA (Los Angeles)值”是对筑路工程中骨料的机械强度的一种度量,结果以“LA”为计量单位。在某种标准物质的定值过程中,准备了大量特定骨料样本,其中一些样本用于确定指定值及其不确定度,28个实验室参加,得到指定值为  $X_{\text{CRM}} = 21.62 \text{ LA}$ ,标准不确定度  $u_{X;\text{CRM}} = 0.26 \text{ LA}$ 。剩余的骨料样本可用于能力验证。

## 5.4 参照值(见 GB/T 15483.1—1999, A.1.1, c)

### 5.4.1 总则

首先准备将要制成标准物质(RM)的测试物料。随机选择一部分样本,在同一实验室中使用合适的测量方法在重复性条件下(如 GB/T 3358.2 中定义)与 CRM 分别测试。对照 CRM 的有证参照值来校准得到测试材料的指定值  $X_{\text{RM}}$ 。

### 5.4.2 指定值的标准不确定度 $u_x$

当测试物料的指定值是由对测试物料和 CRM 的一系列测试得到时,指定值的标准不确定度由测试结果和 CRM 的有证参照值的不确定度得出。如果测试物料和 CRM 不相似(在基质(基体),成分构造和结果的水平等方面),确定指定值的标准不确定度时还应考虑这部分不确定度。

应用该方法得到的指定值,可溯源到 CRM 的有证参照值,其标准不确定度是可计算的,并且可以避免将 CRM 分发给所有参加者的成本支出,因此好于其他方法。但是,该方法需假定测试条件与测试物料没有交互作用。

5.4.3 中的示例说明了当用测试物料与单一 CRM 的直接比对而得到其指定值时,如何计算不确定度的方法。



5.4.3 示例:骨料的 LA 值

5.3 中的示例提到的 CRM 可用于确定另一个相似骨料 RM 的指定值(见式(3))。这个确定过程需要在同一个实验室中使用相同的测量方法以及在重复性条件下对两种骨料样本进行一系列测试。

$$X_{RM} = X_{CRM} + \bar{D} \dots\dots\dots(3)$$

式中:

$X_{CRM}$ ——CRM 的指定值;

$X_{RM}$ ——RM 的指定值;

$D_i$  ——第  $i$  个 RM 样品测量结果(平均值)和第  $i$  个 CRM 样品的测量结果(平均值)之差  
( $x_{i, RM} - x_{i, CRM}$ );

$\bar{D}$  —— $D_i$  的平均值。

RM 指定值的标准不确定度为:

$$u_{X;RM} = \sqrt{u_{X;CRM}^2 + u_D^2} \dots\dots\dots(4)$$

表 1 给出了经过一系列测试可能获得的数据,并给出  $u_D$  的计算方法。

由表中结果可得,

$$X_{RM} = 21.62 + 1.73 = 23.35(LA) \dots\dots\dots(5)$$

和

$$u_{X;RM} = \sqrt{0.26^2 + 0.24^2} = 0.35(LA) \dots\dots\dots(6)$$

其中 0.26 是 CRM 的指定值的标准不确定度(5.3 中的示例给出),0.24 是  $\bar{D}$  的标准不确定度。

表 1 CRM 和 RM 的平均值之差及其标准不确定度的计算

| 样品                                   | RM         |            | CRM        |            | 平均值之差 RM-CRM<br>LA |
|--------------------------------------|------------|------------|------------|------------|--------------------|
|                                      | 测试 1<br>LA | 测试 2<br>LA | 测试 1<br>LA | 测试 2<br>LA |                    |
| 1                                    | 20.5       | 20.5       | 19.0       | 18.0       | 2.00               |
| 2                                    | 21.1       | 20.7       | 19.8       | 19.9       | 1.05               |
| 3                                    | 21.5       | 21.5       | 21.0       | 21.0       | 0.50               |
| 4                                    | 22.3       | 21.7       | 21.0       | 20.8       | 1.10               |
| 5                                    | 22.7       | 22.3       | 20.5       | 21.0       | 1.75               |
| 6                                    | 23.6       | 22.4       | 20.3       | 20.3       | 2.70               |
| 7                                    | 20.9       | 21.2       | 21.5       | 21.8       | -0.60              |
| 8                                    | 21.4       | 21.5       | 21.9       | 21.7       | -0.35              |
| 9                                    | 23.5       | 23.5       | 21.0       | 21.0       | 2.50               |
| 10                                   | 22.3       | 22.9       | 22.0       | 21.3       | 0.95               |
| 11                                   | 23.5       | 24.1       | 20.8       | 20.6       | 3.10               |
| 12                                   | 22.5       | 23.5       | 21.0       | 22.0       | 1.50               |
| 13                                   | 22.5       | 23.5       | 21.0       | 21.0       | 2.00               |
| 14                                   | 23.4       | 22.7       | 22.0       | 22.0       | 1.05               |
| 15                                   | 24.0       | 24.2       | 22.1       | 21.5       | 2.30               |
| 16                                   | 24.5       | 24.4       | 22.3       | 22.5       | 2.05               |
| 17                                   | 24.8       | 24.7       | 22.0       | 21.9       | 2.80               |
| 18                                   | 24.7       | 25.1       | 21.9       | 21.9       | 3.00               |
| 19                                   | 24.9       | 24.4       | 22.4       | 22.6       | 2.15               |
| 20                                   | 27.2       | 27.0       | 24.5       | 23.7       | 3.00               |
| 差的平均值, $\bar{D}$                     |            |            |            |            | 1.73               |
| 标准差                                  |            |            |            |            | 1.07               |
| $\bar{D}$ 的标准不确定度(标准差/ $\sqrt{20}$ ) |            |            |            |            | 0.24               |
| 注:以上是在 LA 检测中,测量骨料的机械强度而获得的数据。       |            |            |            |            |                    |

## 5.5 专家实验室的公议值(见 GB/T 15483.1—1999,A.1.1,d)

### 5.5.1 总则

如参照值方法(5.4),首先需准备分发给参加者的测试物料。随机选取一部分样本,由一组专家实验室进行分析,这组专家实验室可以是一轮能力验证计划的参加者,在这一轮计划完成后确定指定值及其不确定度。指定值  $X$  由这组专家实验室报告结果的稳健平均值得到,具体计算使用附录 C 中的算法 A。

也可使用其他计算方法代替算法 A,只要该方法有可靠的统计学理论基础并在能力验证报告中描述所使用方法即可。

### 5.5.2 指定值的标准不确定度 $u_x$

当  $p$  个专家实验室均报告了测试物料的测量值  $x_i$  及其测量标准不确定度的估计值  $u_i$ ,且指定值  $X$  是由算法 A 计算得到的稳健平均值,则指定值  $X$  的标准不确定度由下式估计:

$$u_x = \frac{1.25}{p} \times \sqrt{\sum_{i=1}^p u_i^2} \dots\dots\dots (7)$$

当专家实验室没有报告标准不确定度,或不确定度未被独立确认(例如,由实验室认可机构确认)时,指定值的标准不确定度应由 5.6 中方法来估计。

注:对于来自正态分布的一个大样本( $p > 10$ ),系数 1.25 是中位数的标准差和算术平均值的标准差之比。对于服从正态分布的数据,由附录 C 中算法 A 得到的稳健平均值的标准差是未知的,但是它会落在算术平均值的标准差和中位数的标准差之间。因而该公式给出了标准不确定度  $u_x$  的一个保守估计。对于  $p < 10$ ,适合的系数将小于 1.25,因而该公式得到的估计将更为保守。

这种方法的局限性在于这组专家实验室的结果中可能存在未知的偏倚,并且所报出的不确定度可能是不可靠的。

### 5.5.3 示例:混凝土的岩相分析

该方法可用于将骨料分发给参加者以确定样本的岩相组成时,对骨料的分类需要一定的技巧和经验,而且没有可供参考的材料,因而这种情况下一组专家实验室的一致意见可能是确定指定值的最好方法。

## 5.6 参加者的公议值(见 GB/T 15483.1—1999,A.1.1,e)

### 5.6.1 总则

该方法是将一轮能力验证计划中所有参加者报告结果的稳健平均值作为该轮计划中所使用测试材料的指定值  $X$ ,计算使用附录 C 中的算法 A。

其他计算方法也可代替算法 A,只要该方法有牢固的统计学基础并在能力验证报告中描述所使用的方法即可。例如,附录 C.1 中按式(C.1)计算稳健平均值,按式(C.2)计算稳健标准差停止,不再进行迭代。

该方法若结合已标准化的测量方法使用,则是非常有用的。

5.6.2 指定值的标准不确定度  $u_x$ 

当指定值为由算法 A 得出的稳健平均值时,指定值  $X$  的标准不确定度由下式估计:

$$u_x = 1.25 \times s^* / \sqrt{p} \quad \dots\dots\dots (8)$$

其中  $s^*$  是使用附录 C 中算法 A 计算得到的稳健标准差。(此处每个参加者得到的“结果”是测试材料的所有观测值的平均值。)

该方法的局限性在于:

- a) 参加者中可能不存在真正的公议值;
- b) 公议值可能会包含测量方法本身的偏倚,但此偏倚不会在如上计算的指定值的标准不确定度中体现出来。

以上情况在痕量成分的测定中是常见的。

## 5.6.3 示例:抗体浓度的测定

表 2 给出了一轮能力验证计划中对三种特定过敏原免疫球蛋白 E 抗体进行测定的数据。图 2 为这些数据的茎叶图。

应用算法 A,先将数据按升序排列,计算均值和标准差的初始稳健估计值,再用迭代法计算。表 3 给出了表 2 中特定过敏原免疫球蛋白 E 抗体 d1 测量结果的计算过程。

算法 A 的计算过程如下。

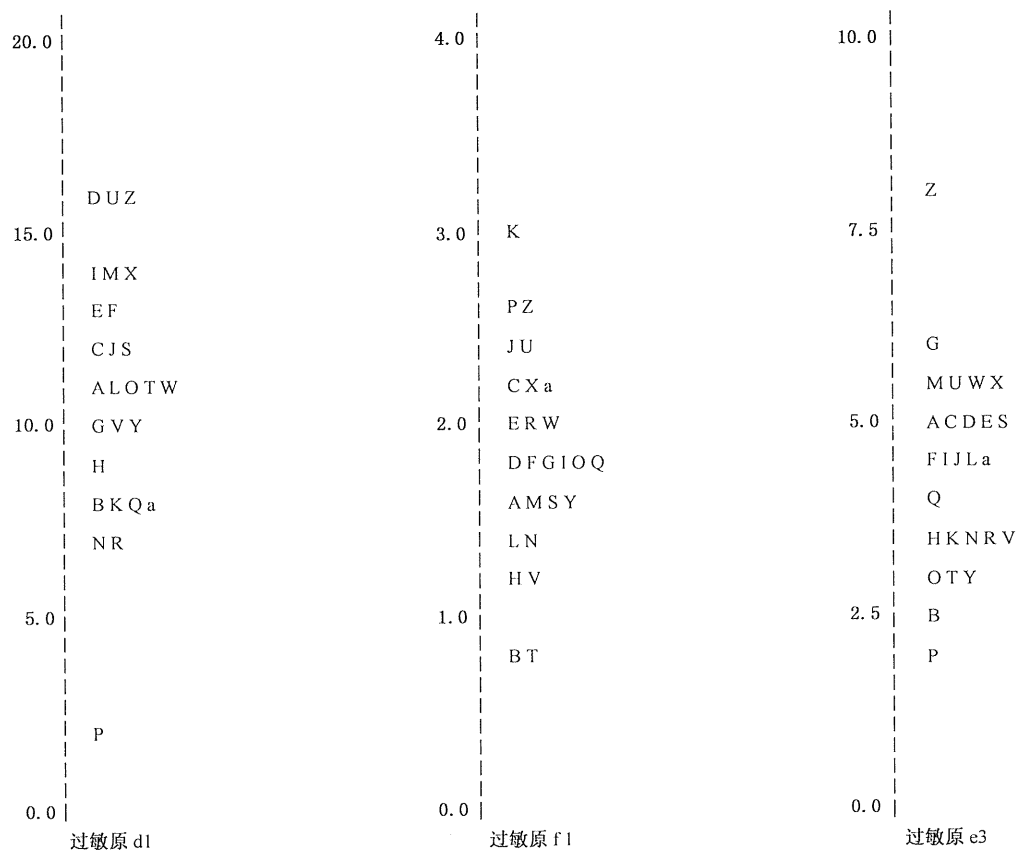
- a) 步骤 1:将数据按升序排在一列中,如表 3 中第 0 步迭代。计算其均值和标准差(表 3 中 10.91 和 3.13)。用 C.1 中的公式计算稳健平均值和稳健标准差的初始值(表 3 中的 10.85 和 3.53)。
- b) 步骤 2:将数据复制到第二列中,如表 3 第 1 步迭代所示。用 C.1 中的公式由稳健平均值和稳健标准差的初始值计算截止值,(表 3 中的 5.56 和 16.15)。将超出截止值范围的数值用截止值代替(2.18 替换为 5.56,16.30 替换为 16.15)。计算新的均值和标准差(表 3 中 11.03 和 2.81)。根据附录 C 给出的公式,稳健平均值与此时的均值(11.03)相同,将标准差乘以 1.134 得到稳健标准差(3.19)。
- c) 步骤 3:继续使用数据表,不必将数据写入新列中。第三列顶部的截止值需改变,即用此列底部的稳健平均值和稳健标准差计算得新的截止值。由第二步迭代可得截止值(6.24 和 15.82),如表 3 所示。继续将超出截止值范围的数值替换为截止值,直到迭代终止。当数据进行替换时,数据表将更新均值、标准差和截止值,但这些数值的变化将逐渐变小直到不再显著变化为止。

其他两种特定过敏原免疫球蛋白 E 抗体的稳健平均值和稳健标准差可用相似的方法进行计算。

值得注意的是,实验室 P 对 d1 的测试结果没有按照格拉布斯(Grubbs)检验归为离群值或岐离值,示例中若根据 GB/T 6379.2 对于离群值的判定,则应由表 3 中第 0 次迭代的数据计算出均值和标准差。应用稳健方法,实验室 P 得到的较小结果,以及实验室 D、U 和 Z 得到的较大结果对稳健估计值没有影响。从条形图(见 8.3 图 9)中可以得出某些实验室结果可能需要进一步调查。例如,图 9 中,实验室 Z 在所有 3 个水平上具有最大的正的  $z$  值,说明该实验室可能存在较大的正偏倚。

表 2 三种特定过敏原免疫球蛋白 E 抗体(d1,f1 和 e3)的浓度——来自  $p=27$  个实验室报告的数据

| 实验室                                            | 浓度         |            |            |
|------------------------------------------------|------------|------------|------------|
|                                                | d1<br>kU/L | f1<br>kU/L | e3<br>kU/L |
| A                                              | 11.30      | 1.69       | 5.02       |
| B                                              | 8.29       | 0.74       | 2.52       |
| C                                              | 11.90      | 2.23       | 5.15       |
| D                                              | 15.60      | 1.76       | 5.15       |
| E                                              | 13.40      | 1.91       | 4.84       |
| F                                              | 12.50      | 1.71       | 4.54       |
| G                                              | 10.40      | 1.88       | 5.94       |
| H                                              | 9.38       | 1.14       | 3.50       |
| I                                              | 14.20      | 1.74       | 4.48       |
| J                                              | 12.10      | 2.39       | 4.75       |
| K                                              | 8.10       | 3.10       | 3.70       |
| L                                              | 10.80      | 1.39       | 4.70       |
| M                                              | 13.80      | 1.52       | 5.59       |
| N                                              | 7.00       | 1.50       | 3.40       |
| O                                              | 10.85      | 1.80       | 2.80       |
| P                                              | 2.18       | 2.52       | 1.88       |
| Q                                              | 8.39       | 1.83       | 3.80       |
| R                                              | 6.95       | 1.92       | 3.52       |
| S                                              | 11.80      | 1.58       | 4.86       |
| T                                              | 10.90      | 0.80       | 2.80       |
| U                                              | 16.30      | 2.39       | 5.60       |
| V                                              | 9.71       | 1.21       | 3.33       |
| W                                              | 10.50      | 1.93       | 5.35       |
| X                                              | 13.60      | 2.23       | 5.53       |
| Y                                              | 10.10      | 1.63       | 3.18       |
| Z                                              | 16.07      | 2.69       | 8.22       |
| a                                              | 8.47       | 2.16       | 4.64       |
| p                                              | 27         | 27         | 27         |
| 稳健均值 $\bar{x}^*$                               | 11.03      | 1.83       | 4.35       |
| 稳健标准差 $s^*$                                    | 3.04       | 0.50       | 1.25       |
| 注：表中数据为每升(L)样品中千单位(kU)抗体数，其中单位(U)由国际标准物质的浓度定义。 |            |            |            |



注 1：图中数据为每升(L)样品中千单位(kU)抗体数，其中一单位(U)由国际标准物质的浓度定义。  
注 2：表 2 中的数值为手工计算到小数点后两位。

图 2 三种特定过敏原免疫球蛋白 E 抗体的浓度——数据的茎叶图(由表 2 中数据)

表 3 三种特定过敏原免疫球蛋白 E 抗体的浓度——抗体 d1 的稳健平均值和稳健标准差的计算

| 迭代步骤            | 0     | 1     | 2     | 3     | 4     | 5     |
|-----------------|-------|-------|-------|-------|-------|-------|
| $\delta=1.5s^*$ | —     | 5.30  | 4.79  | 4.62  | 4.58  | 4.56  |
| $x^*-\delta$    | —     | 5.56  | 6.24  | 6.41  | 6.45  | 6.47  |
| $x^*+\delta$    | —     | 16.15 | 15.82 | 15.65 | 15.61 | 15.59 |
| P               | 2.18  | 5.56  | 6.24  | 6.41  | 6.45  | 6.47  |
| R               | 6.95  | 6.95  | 6.95  | 6.95  | 6.95  | 6.95  |
| N               | 7.00  | 7.00  | 7.00  | 7.00  | 7.00  | 7.00  |
| K               | 8.10  | 8.10  | 8.10  | 8.10  | 8.10  | 8.10  |
| B               | 8.29  | 8.29  | 8.29  | 8.29  | 8.29  | 8.29  |
| Q               | 8.39  | 8.39  | 8.39  | 8.39  | 8.39  | 8.39  |
| a               | 8.47  | 8.47  | 8.47  | 8.47  | 8.47  | 8.47  |
| H               | 9.38  | 9.38  | 9.38  | 9.38  | 9.38  | 9.38  |
| V               | 9.71  | 9.71  | 9.71  | 9.71  | 9.71  | 9.71  |
| Y               | 10.10 | 10.10 | 10.10 | 10.10 | 10.10 | 10.10 |

表 3 (续)

| 迭代步骤     | 0     | 1     | 2     | 3     | 4     | 5     |
|----------|-------|-------|-------|-------|-------|-------|
| G        | 10.40 | 10.40 | 10.40 | 10.40 | 10.40 | 10.40 |
| W        | 10.50 | 10.50 | 10.50 | 10.50 | 10.50 | 10.50 |
| L        | 10.80 | 10.80 | 10.80 | 10.80 | 10.80 | 10.80 |
| O        | 10.85 | 10.85 | 10.85 | 10.85 | 10.85 | 10.85 |
| T        | 10.90 | 10.90 | 10.90 | 10.90 | 10.90 | 10.90 |
| A        | 11.30 | 11.30 | 11.30 | 11.30 | 11.30 | 11.30 |
| S        | 11.80 | 11.80 | 11.80 | 11.80 | 11.80 | 11.80 |
| C        | 11.90 | 11.90 | 11.90 | 11.90 | 11.90 | 11.90 |
| J        | 12.10 | 12.10 | 12.10 | 12.10 | 12.10 | 12.10 |
| F        | 12.50 | 12.50 | 12.50 | 12.50 | 12.50 | 12.50 |
| E        | 13.40 | 13.40 | 13.40 | 13.40 | 13.40 | 13.40 |
| X        | 13.60 | 13.60 | 13.60 | 13.60 | 13.60 | 13.60 |
| M        | 13.80 | 13.80 | 13.80 | 13.80 | 13.80 | 13.80 |
| I        | 14.20 | 14.20 | 14.20 | 14.20 | 14.20 | 14.20 |
| D        | 15.60 | 15.60 | 15.60 | 15.60 | 15.60 | 15.59 |
| Z        | 16.07 | 16.07 | 15.82 | 15.65 | 15.61 | 15.59 |
| U        | 16.30 | 16.15 | 15.82 | 15.65 | 16.61 | 15.59 |
| 平均值      | 10.91 | 11.03 | 11.03 | 11.03 | 11.03 | 11.03 |
| 标准差      | 3.13  | 2.81  | 2.72  | 2.69  | 2.68  | 2.68  |
| 新的 $x^*$ | 10.85 | 11.03 | 11.03 | 11.03 | 11.03 | 11.03 |
| 新的 $s^*$ | 3.53  | 3.19  | 3.08  | 3.05  | 3.04  | 3.04  |

注：表中数据为每升(L)样品中千单位(kU)抗体数，其中单位(U)由国际标准物质的浓度定义。

## 5.7 指定值的比对

当使用 5.2 至 5.4 中方法计算指定值  $X$  时，在每轮能力验证计划后，应将每轮计划的结果得到的稳健平均值  $x^*$  与指定值比较。当使用 5.5 和 5.6 中方法计算指定值时，若有可能，该指定值应与一个具备专业能力的实验室得到的参照值作比较。差值  $x^* - X$  的标准不确定度由下式估计：

$$\sqrt{\frac{(1.25s^*)^2}{p} + u_x^2} \dots\dots\dots (9)$$

其中， $s^*$  为稳健标准差； $p$  为实验室数。

若  $|x^* - X|$  比其不确定度的二倍还要大，则需查找原因。可能的原因如下：

- 测量方法偏倚；
- 实验室偏倚；
- 没有意识到使用 5.2 中方法的局限性；
- 当使用“专家实验室”的公议值作为指定值时，专家实验室存在偏倚；
- 当使用参加者的公议值作为指定值时，使用的测量方法存在偏倚，或某些实验室存在偏倚。

## 5.8 缺失值

当能力验证中重复测试次数  $n$  不少于 2 时,则分析结果时应首先计算各实验室结果的均值和标准差。这个均值可用于以下情形:例如,在第 7 章中性能统计量的计算,8.2 和 8.3 中绘制直方图和条形图。标准差可用于绘制如 8.4 和 8.6 中考察测量重复性的图。

尽管所有参加者预期进行相同次数  $n$  的重复测试,但他们报告的测试结果可能有缺失,例如,某些测试被毁坏或不能重复。当以上情况发生时,推荐按以下步骤进行。

若实验室报告了至少  $0.59n$  次重复测试结果,仍可用这些数据计算均值和标准差,并看作实验室已报告了  $n$  个测试值。但报告中须说明实验室所做重复测试的次数。

若实验室报告的重复测试次数小于  $0.59n$ ,则不应将该测量结果用于影响其他实验室的统计量计算中。例如,其结果不应包含在如 5.6 中指定值的计算中,或如 6.6 中能力评定标准差的计算中。这些测量结果可用于如第 7 章其自身性能统计量的计算中,或者用于第 8 章的图示中,但是报告中应声明实验室所作的重复测试次数,并说明测试次数少于计划中预期的测试次数。

注:系数定为 0.59 是因为: $n$  次重复测试均值的标准差是  $\sigma_r/\sqrt{n}$ 。由于重复测试次数减少则标准差增大,因而当实际重复测试次数从  $n$  减少到  $0.59n$  时,标准差增大到原来的 1.3 倍。此时可认为达到标准差波动的一个可接受的极限值。由于  $0.59n$  的限制,这将阻止标准差大于其 1.3 倍。该规则一定程度上是人为规定的,因此如有必要,协调者可与计划参加者商议修改规则。

## 6 确定能力评定标准差(见 GB/T 15483.1—1999, A.2.1.3)

### 6.1 方法选择

6.2 至 6.6 中描述了五种确定能力评定标准差  $\hat{\sigma}$  的方法。协调者应负责对方法的选择,方法选择过程中应与计划的参加者和相关认可机构进行协商,并考虑相关规定。当参加能力验证计划的实验室很少时,6.6 中方法是不适用的。协调者还应负责确定标准差  $\hat{\sigma}$ ,并报告获得标准差的具体步骤。

注:GB/T 15483.1 使用符号  $s$  表示能力评定标准差。这与  $s$  的通常用法是一致的, $s$  一般用于表示样本标准差。

本标准中,能力评定标准差有时也可由其他方法得出,为不引起混淆,此处选择  $\hat{\sigma}$  表示能力评定标准差。

### 6.2 由规定值确定

#### 6.2.1 总则

能力评定标准差可根据某个特定的数据解释工作的需要设定为一规定值,或由法规要求确定。

这种方法的优点在于能力评定标准差与测量方法的“适用性”声明有直接关联。

#### 6.2.2 示例:坚果,坚果制品,干无花果和干无花果制品中黄曲霉素含量的测定

法规规定,当黄曲霉素的法定上限值为  $10 \mu\text{g}/\text{kg}$  时,测定黄曲霉素方法的再现性变异系数应不大于 50%。因此,当用于能力验证的实验材料中黄曲霉素含量为  $10 \mu\text{g}/\text{kg}$  时,由法规可知测量结果的再现性标准差应不大于  $5 \mu\text{g}/\text{kg}$ 。这种情况下,设定能力评定标准差为  $5 \mu\text{g}/\text{kg}$  是适合的。

### 6.3 由经验预期值确定

#### 6.3.1 总则

能力评定标准差可设定为某一符合实验室能力水平的值,它是协调者和能力验证计划成员希望实验室可以达到的预期值。

因而,由此设定实验能力评定标准差就等价于关于测量方法的“适用性”声明。

当能力评定标准差  $\hat{\sigma}$  是规定值或由经验预期值得到时,则所取数值与测量方法的再现性可能是没有实际联系的。只要该方法的重复性和再现性是已知的,以下方法可用于检查所选取的  $\hat{\sigma}$  是否合理。已知

$\sigma_R$  为再现性标准差,

$\sigma_r$  为重复性标准差,

计算实验室间标准差:

$$\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2}$$

将  $\sigma_L$ 、 $\sigma_r$  和选定的  $\hat{\sigma}$  的值代入式(10)中,可得系数  $\phi$ ,

$$\hat{\sigma} = \sqrt{(\phi \times \sigma_L)^2 + (\sigma_r^2/n)} \quad \dots\dots\dots(10)$$

其中  $n$  是能力验证中各实验室将实施的重复测试次数。

当所得  $\phi$  很小(如  $\phi < 0.5$ )时,表明实际上实验室无法达到与选定  $\hat{\sigma}$  对应的再现性水平。

### 6.3.2 示例 1:人血清中血糖含量的测定

假定医学实验室能够测定的血糖水平值在指定值的  $\pm 10\%$  以内是可接受的,而对于极低的血糖浓度(低于 60 mg/dL)的允许差为  $\pm 6.0$  mg/dL。由此可计算能力评定标准差:

a) 对于指定值  $X$  低于 60 mg/dL:

$$\hat{\sigma} = 6.0/3.0 = 2.0 \text{ mg/dL};$$

b) 对于指定值  $X$  高于 60 mg/dL:

$$\hat{\sigma} = 0.1X/3.0 = 0.033X \text{ mg/dL}。$$

此处的系数 3.0 对应于用来解释  $z$  值(见 7.4)的临界值 3.0。

### 6.3.3 示例 2:测定硬化混凝土中水泥含量

混凝土中水泥含量通常由每立方米混凝土中水泥的质量来度量(如, kg/m<sup>3</sup>)。实际上,混凝土的生产按其质量不同而分为不同等级,以 25 kg/m<sup>3</sup> 为间隔定级,实验室需能正确测定混凝土等级。由于以上原因, $\hat{\sigma}$  的取值不应大于 25 kg/m<sup>3</sup> 的一半。对于一种水泥平均含量为 260 kg/m<sup>3</sup> 的混凝土,通过精密度试验结果可知: $\sigma_R = 23.2$  kg/m<sup>3</sup>,  $\sigma_r = 14.3$  kg/m<sup>3</sup>。

因此

$$\sigma_L = \sqrt{23.2^2 - 14.3^2} = 18.3(\text{kg/m}^3) \quad \dots\dots\dots(11)$$

若  $n=2$ ,将  $\sigma_L = 18.3$  kg/m<sup>3</sup>、 $\sigma_r = 14.3$  kg/m<sup>3</sup> 和  $\hat{\sigma} = 12.5$  kg/m<sup>3</sup> 代入式(10)可得:

$$12.5 = \sqrt{(18.3\phi)^2 + (14.3^2/2)} \quad \dots\dots\dots(12)$$

由上式可得  $\phi = 0.40$ 。因此,若取  $\hat{\sigma} = 12.5$  kg/m<sup>3</sup>,则能力验证中要求的实验室间标准差是精密密度试验中所得实验室间标准差的 0.4 倍,显然这是不可实现的。

## 6.4 由一般模型确定

### 6.4.1 总则

能力评定标准差可由测量方法再现性的一般模型得出。

该方法的缺点在于,特定测量方法真实的再现性可能与一般模型给出的数值有显著差别,这是因为一般模型假定再现性只依赖于被测量的水平,而不依赖于被测量本身、测量过程或者样本量。

### 6.4.2 示例:Horwitz 曲线

Horwitz<sup>[3]</sup>给出了化学分析方法再现性的一个一般模型,这个方法可得到以下再现性标准差的表



达式:

$$\sigma_R = 0.02c^{0.8495} \dots\dots\dots (13)$$

其中  $c$  是以百分数表示的待测化学成分的浓度(质量分数)。

## 6.5 由精密度试验结果确定

### 6.5.1 总则

当能力验证计划中使用已标准化的测量方法,且该方法的重复性和再现性可获得时,能力评定标准差  $\hat{\sigma}$  可由这些信息进行如下计算。已知

$\sigma_R$  为再现性标准差,

$\sigma_r$  为重复性标准差,

实验室间标准差计算公式如下:

$$\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2} \dots\dots\dots (14)$$

则计算能力评定标准差的公式如下:

$$\hat{\sigma} = \sqrt{\sigma_L^2 + (\sigma_r^2/n)} \dots\dots\dots (15)$$

其中  $n$  是一轮能力验证计划中各实验室的重复测试次数。

当重复性标准差和再现性标准差依赖于测试结果的平均值时,它们之间的函数关系可以用 GB/T 6379.2 中的方法确定,然后利用这些关系对能力验证的指定值计算重复性标准差和再现性标准差。

### 6.5.2 示例:测定硬化混凝土中水泥含量

沿用 6.3 示例中的数据,式(15)给出了能力评定标准差如下:

$$\hat{\sigma} = \sqrt{18.3^2 + (14.3^2/2)} = 20.9(\text{kg/m}^3) \dots\dots\dots (16)$$

其中假定能力验证计划中重复测试次数  $n=2$ 。

## 6.6 由一轮能力验证计划所得数据确定

### 6.6.1 总则

该方法中,用于评估一轮能力验证计划中参加者能力的标准差  $\hat{\sigma}$  由该轮计划参加者报告的结果得出。标准差应是所有参加者报告结果的稳健标准差,用附录 C 中算法 A 计算。参加者报告的结果应是此轮计划中实验室  $n$  次重复测试的平均值。

也可用有合理统计学理论基础的其他计算方法代替算法 A,此时应在报告中描述该方法。

该方法的缺点在于,由于  $\hat{\sigma}$  的值可能在每轮计划都会有显著的变化,因而利用实验室  $z$  值来寻找几轮计划中可能的趋势时会有一定困难。但这个缺点可通过制定能力验证计划时,规定  $\hat{\sigma}$  为若干轮计划中所得标准差的稳健联合值(由附录 C 中算法 S 计算)来克服。

### 6.6.2 示例:测定抗体浓度

表 2 和表 3 给出了该方法的一个示例。

## 6.7 能力验证得到的精密度与测量方法已知精密度的比较

为了检验参加者能力,并衡量能力验证计划对参加者的益处,推荐协调者采用以下步骤。使用 GB/T 6379.5 中的稳健方法,通过各轮能力验证计划的结果来估计测量方法的重复性标准差和再现性

标准差。画出这些估计值的时间序列图,如有可能,图中也应标出精密度试验所得的重复性标准差和再现性标准差。

之后协调者应仔细检查这些图。若能力验证中所得精密度与精密度试验中所得数值相差两倍或更多,协调者需调查其原因。如果能力验证计划得到的测量方法精密度没有随时间推移而改善,则建议考虑以下几点:

- 参加实验室未调查发出行动信号或警戒信号的原因,或者没有正确执行纠正措施;
- 参加实验室不能查出发出行动信号或警戒信号的原因;
- 所采用方法在测量过程中处于统计受控状态,可基于测得的数据得出可靠的结论。

## 7 性能统计量的计算

### 7.1 实验室偏倚的估计 (见 GB/T 15483.1—1999, A.2.1.4, a)

#### 7.1.1 总则

设  $x$  表示一轮能力验证计划中参加者对测试材料的某种特性报告的结果。则该特性的实验室偏倚估计值  $D$  由下式计算:

$$D = x - X \quad \dots\dots\dots (17)$$

其中  $X$  为指定值。

性能统计量中不推荐使用  $|D|$  或  $D^2$ , 这样会掩盖偏倚的符号。

#### 7.1.2 实验室偏倚的解释

当参加者所报结果显示实验室偏倚大于  $3.0\hat{\sigma}$  或小于  $-3.0\hat{\sigma}$  时,则发出“行动信号”。同样的,当实验室偏倚大于  $2.0\hat{\sigma}$  或小于  $-2.0\hat{\sigma}$  则发出“警戒信号”。一轮计划中的一个“行动信号”,或连续两轮计划中出现“警戒信号”,应认为出现异常数据,需进行调查。该准则等同于 7.4 中根据  $z$  值给出行动信号和警戒信号的准则。

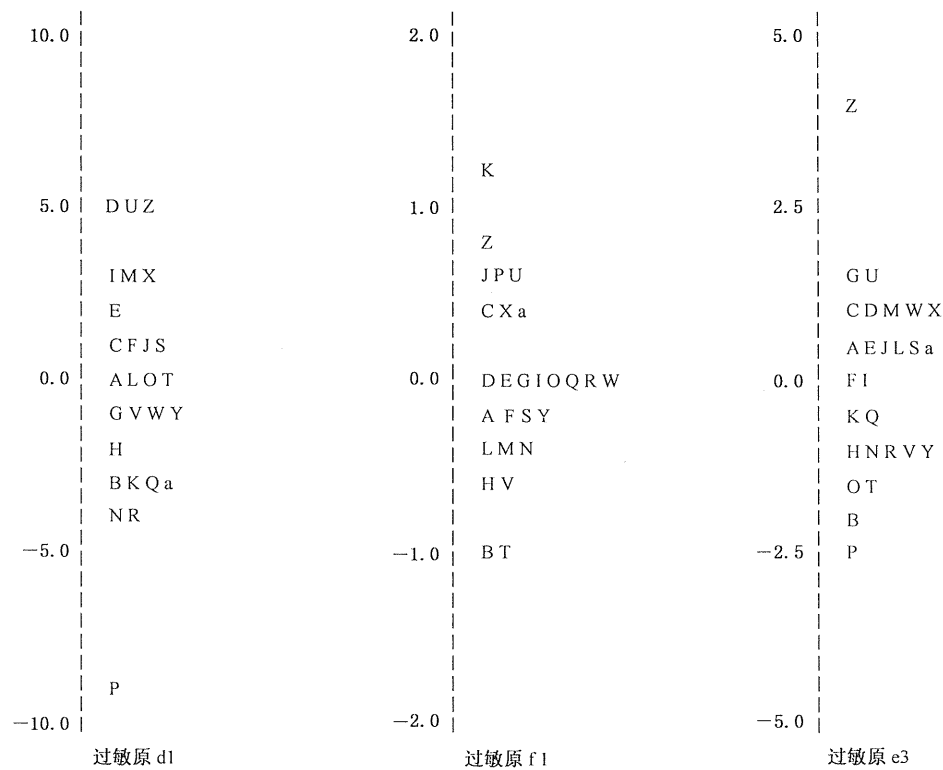
这里系数为 2.0 和 3.0 (此后也出现在相似的准则中) 的原因如下: 若  $x$  来自于服从正态分布的总体, 且  $X$  和  $\hat{\sigma}$  是总体均值和标准差的优良估计值, 那么  $D$  值也近似服从正态分布, 其均值为零, 标准差为  $\hat{\sigma}$ 。在这些条件下, 实验室偏倚的估计值落在  $-3.0\hat{\sigma} < D < 3.0\hat{\sigma}$  的区间之外的概率只有 0.3%, 而落在  $-2.0\hat{\sigma} < D < 2.0\hat{\sigma}$  的区间之外的概率则有 5.0%。正因为行动信号的发生概率很低, 当没有真正的问题存在时行动信号是很难出现的, 所以若出现了行动信号则认为有异常情况出现是合理的。

当能力评定标准差由 6.2、6.3 或 6.4 中方法确定时, 它可能与再现性标准差有显著不同, 此时概率 0.3% 和 5.0% 就不再适用。

当能力评定标准差由 6.2 或 6.3 中方法确定时, 可根据实际情况选择适当的性能统计量, 使之可以直接与规定值或经验预期值作比较。例如, 6.3 的示例中, 样本中血糖含量指定值的误差限为  $\pm 10\%$ , 能力评定标准差为 3.33%, 再按 7.4 中方法计算  $z$  值。同样地, 也可计算实验室偏倚的百分相对差 (见 7.2), 然后直接与目标值 10% 比较。

#### 7.1.3 示例: 抗体浓度的测定

表 4 给出用以上方法处理表 2 数据后所得结果, 图 3 给出了实验室偏倚估计值的茎叶图。比较图 2 和图 3 可知, 实验室偏倚与原始数据具有相同的分布 (不考虑修约的影响), 但是实验室偏倚以零点为中心。



注：图中数据为每升(L)样品中千单位(kU)抗体数,其中单位(U)由国际标准物质的浓度定义。

图 3 三种特定过敏原免疫球蛋白 E 抗体的浓度——实验室偏倚估计值的茎叶图(由表 4 数据)

表 4 三种特定过敏原免疫球蛋白 E 抗体的浓度(d1,f1 和 e3)——实验室偏倚的估计值

| 实验室 | 浓度         |            |            |
|-----|------------|------------|------------|
|     | d1<br>kU/L | f1<br>kU/L | e3<br>kU/L |
| A   | 0.27       | -0.14      | 0.67       |
| B   | -2.74      | -1.09W     | -1.83      |
| C   | 0.87       | 0.40       | 0.80       |
| D   | 4.57       | -0.07      | 0.80       |
| E   | 2.37       | 0.08       | 0.49       |
| F   | 1.47       | -0.12      | 0.19       |
| G   | -0.63      | 0.05       | 1.59       |
| H   | -1.65      | -0.69      | -0.85      |
| I   | 3.17       | -0.09      | 0.13       |
| J   | 1.07       | 0.56       | 0.40       |
| K   | -2.93      | 1.27W      | -0.65      |
| L   | -0.23      | -0.44      | 0.35       |
| M   | 2.77       | -0.31      | 1.24       |
| N   | -4.03      | -0.33      | -0.95      |

表 4 (续)

| 实验室                | 浓度         |            |            |
|--------------------|------------|------------|------------|
|                    | d1<br>kU/L | f1<br>kU/L | e3<br>kU/L |
| O                  | −0.18      | −0.03      | −1.55      |
| P                  | −8.85W     | 0.69       | −2.47      |
| Q                  | −2.64      | 0.00       | −0.55      |
| R                  | −4.08      | 0.09       | −0.83      |
| S                  | 0.77       | −0.25      | 0.51       |
| T                  | −0.13      | −1.03W     | −1.55      |
| U                  | 5.27       | 0.56       | 1.25       |
| V                  | −1.32      | −0.62      | −1.02      |
| W                  | −0.53      | 0.10       | 1.00       |
| X                  | 2.57       | 0.40       | 1.18       |
| Y                  | −0.93      | −0.20      | −1.17      |
| Z                  | 5.04       | 0.86       | 3.87A      |
| a                  | −2.56      | 0.33       | 0.29       |
| $3.0\hat{\sigma}$  | 9.12       | 1.50       | 3.75       |
| $2.0\hat{\sigma}$  | 6.08       | 1.00       | 2.50       |
| $−2.0\hat{\sigma}$ | −6.08      | −1.00      | −2.50      |
| $−3.0\hat{\sigma}$ | −9.12      | −1.50      | −3.75      |

注 1: 浓度数据中,A=行动信号,W=警戒信号。  
 注 2: 表中的实验室偏倚数据是根据表 2 中的数据计算,将表 2 的稳健平均值作为三个水平的指定值。上表底部给出了行动限和警戒限,其中将表 2 的稳健标准差作为能力评定标准差。  
 注 3: 数据为每升(L)样品中千单位(kU)抗体数,其中单位(U)由国际标准物质的浓度定义。

## 7.2 百分相对差(见 GB/T 15483.1—1999,A.2.1.4,b)

### 7.2.1 总则

沿用 7.1 中符号,百分相对差由下式定义:

$$D\% = 100(x - X)/X \quad \dots\dots\dots(18)$$

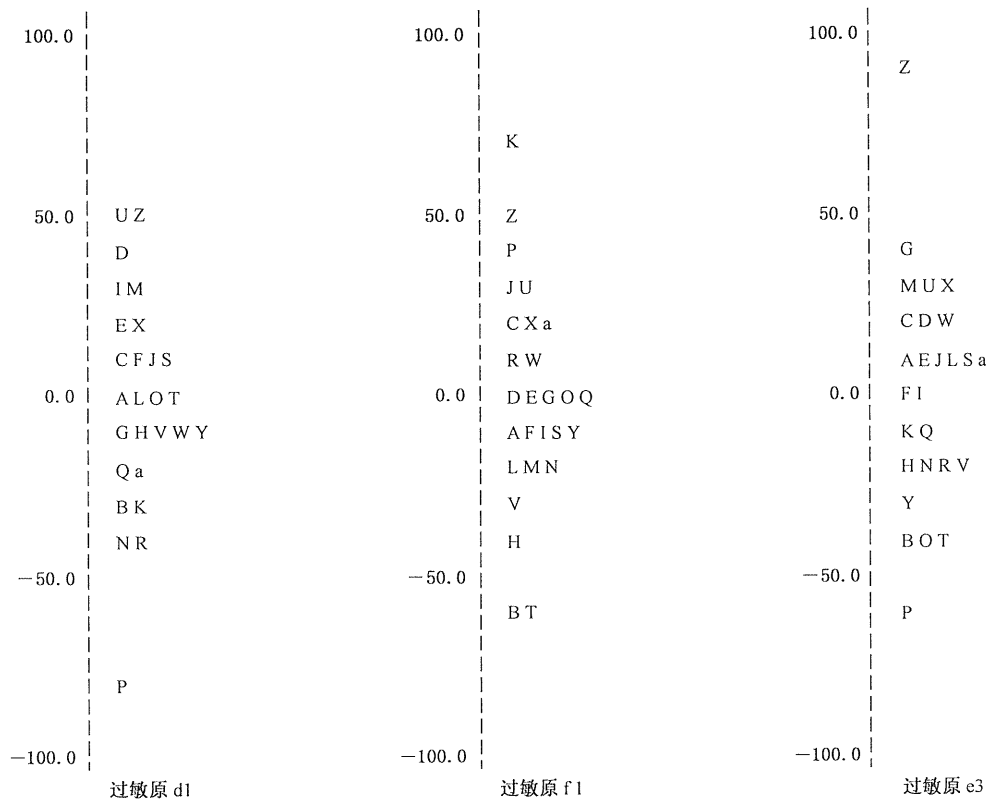
### 7.2.2 百分相对差的解释

百分相对差应与 7.1 中实验室偏倚有等价的解释,即当实验室报告结果的百分相对差大于  $300\hat{\sigma}/X$  或小于  $-300\hat{\sigma}/X$  时,应认为产生了一个“行动信号”。同样的,百分相对差大于  $200\hat{\sigma}/X$  或小于  $-200\hat{\sigma}/X$  可认为产生一个“警戒信号”。一个“行动信号”或连续两轮计划中的“警戒信号”应认为出现了异常。

### 7.2.3 示例:抗体浓度的测定

表 5 给出了由该方法处理表 2 数据所得的结果,图 4 显示了百分相对差的茎叶图。比较图 2 和图 4 可知,它与原始数据(不考虑修约的影响)具有相同的分布,但是百分相对差是以零点为中心,这与实验

室偏倚的情况相似。



注：图中数据为每升(L)样品中千单位(kU)抗体数,其中单位(U)由国际标准物质的浓度定义。

图 4 三种特定过敏原免疫球蛋白 E 抗体的浓度——百分相对差的茎叶图(由表 5 数据)

表 5 三种特定过敏原免疫球蛋白 E 抗体的浓度(d1,f1 和 e3)——百分相对差

| 实验室 | 浓度      |         |         |
|-----|---------|---------|---------|
|     | d1<br>% | f1<br>% | e3<br>% |
| A   | 2       | —8      | 15      |
| B   | —25     | —60W    | —42     |
| C   | 8       | 22      | 18      |
| D   | 41      | —4      | 18      |
| E   | 21      | 4       | 11      |
| F   | 13      | —7      | 4       |
| G   | —6      | 3       | 37      |
| H   | —15     | —38     | —20     |
| I   | 29      | —5      | 3       |
| J   | 10      | 31      | 9       |
| K   | —27     | 69W     | —15     |
| L   | —2      | —24     | 8       |

表 5 (续)

| 实验室                                                                                             | 浓度      |         |         |
|-------------------------------------------------------------------------------------------------|---------|---------|---------|
|                                                                                                 | d1<br>% | f1<br>% | e3<br>% |
| M                                                                                               | 25      | -17     | 29      |
| N                                                                                               | -37     | -18     | -22     |
| O                                                                                               | -2      | -2      | -36     |
| P                                                                                               | -80W    | 38      | -57     |
| Q                                                                                               | -24     | 0       | -13     |
| R                                                                                               | -37     | 5       | -19     |
| S                                                                                               | 7       | -14     | 12      |
| T                                                                                               | -1      | -56     | -36     |
| U                                                                                               | 48      | 31      | 29      |
| V                                                                                               | -12     | -34     | -23     |
| W                                                                                               | -5      | 5       | 23      |
| X                                                                                               | 23      | 22      | 27      |
| Y                                                                                               | -8      | -11     | -27     |
| Z                                                                                               | 46      | 47      | 89A     |
| a                                                                                               | -23     | 18      | 7       |
| $300\hat{\sigma}/X$                                                                             | 82.7    | 82.0    | 86.2    |
| $200\hat{\sigma}/X$                                                                             | 55.1    | 54.6    | 57.5    |
| $-200\hat{\sigma}/X$                                                                            | -55.1   | -54.6   | -57.5   |
| $-300\hat{\sigma}/X$                                                                            | -82.7   | -82.0   | -86.2   |
| 注 1: 浓度数据中,A=行动信号,W=警戒信号。                                                                       |         |         |         |
| 注 2: 表中的百分相对差是由表 2 中的数据计算得到,其中将表 2 数据的稳健均值作为三个水平的指定值。表的底部给出了行动限和警戒限,其中将表 2 的稳健标准差作为能力评定标准差进行计算。 |         |         |         |

### 7.3 秩与秩百分数(见 GB/T 15483.1—1999,A.2.1.4,c)

#### 7.3.1 总则

根据  $p$  个实验室在一轮能力验证计划的结果,指定报告结果最小的实验室秩为 1,结果第二小的实验室秩为 2,以此类推,直到报告最大结果的实验室秩为  $p$ ,由此来确定实验室的秩。若有两个或更多的结果相同,则将这几个实验室指定为相同的平均秩。例如,表 2 中实验室 C 和 X 报告 f1 的浓度均为 2.23,他们的秩为 21 和 22,所以表 6 中将他们的秩指定为 21.5。若实验中包括了几个被测量,则对每个被测量分别指定秩。

若秩表示为  $i=1,2,\dots,p$ ,则秩百分数为  $100(i-0.5)/p\%$ 。表 6 为计算秩百分数的示例。

#### 7.3.2 秩和秩百分数的解释

秩或秩百分数的解释不需假设数据服从某种特定的概率分布,且其方差不需用到指定值或能力评

定标准差。因此,秩和秩百分数提供了一种确定报告结果最极端的实验室的简单方法。在最初几轮能力验证计划中,这个方法能判定最有可能改进实验能力的实验室,这是非常有用的。然而,应注意到GB/T 15483.1—1999 中 6.6.5 的警告:“在能力验证中,不推荐对实验室按秩的方式出具报告。因此,为避免引起误导和造成误解,对秩应当极其慎重”。

表 6 三种特定过敏原免疫球蛋白 E 抗体的浓度(d1,f1 和 e3)——秩和秩百分数

| 实验室 | 秩  |      |      | 秩百分数    |         |         |
|-----|----|------|------|---------|---------|---------|
|     | d1 | f1   | e3   | d1<br>% | f1<br>% | e3<br>% |
| A   | 16 | 10   | 19   | 57      | 35      | 69      |
| B   | 5  | 1    | 2    | 17      | 2       | 6       |
| C   | 18 | 21.5 | 20.5 | 65      | 78      | 74      |
| D   | 25 | 13   | 20.5 | 91      | 46      | 74      |
| E   | 21 | 17   | 17   | 76      | 61      | 61      |
| F   | 20 | 11   | 13   | 72      | 39      | 46      |
| G   | 11 | 16   | 26   | 39      | 57      | 94      |
| H   | 8  | 3    | 8    | 28      | 9       | 28      |
| I   | 24 | 12   | 12   | 87      | 43      | 43      |
| J   | 19 | 23.5 | 16   | 60      | 85      | 57      |
| K   | 4  | 27   | 10   | 13      | 98      | 35      |
| L   | 13 | 5    | 15   | 46      | 17      | 54      |
| M   | 23 | 7    | 24   | 83      | 24      | 87      |
| N   | 3  | 6    | 7    | 9       | 20      | 24      |
| O   | 14 | 14   | 3.5  | 50      | 50      | 11      |
| P   | 1  | 25   | 1    | 2       | 91      | 2       |
| Q   | 6  | 15   | 11   | 20      | 54      | 39      |
| R   | 2  | 18   | 9    | 6       | 65      | 31      |
| S   | 17 | 8    | 18   | 61      | 28      | 65      |
| T   | 15 | 2    | 3.5  | 54      | 6       | 11      |
| U   | 27 | 23.5 | 25   | 98      | 85      | 91      |
| V   | 9  | 4    | 6    | 31      | 13      | 20      |
| W   | 12 | 19   | 22   | 43      | 69      | 80      |
| X   | 22 | 21.5 | 23   | 80      | 78      | 83      |
| Y   | 10 | 9    | 5    | 35      | 31      | 17      |
| Z   | 26 | 26   | 27   | 94      | 94      | 98      |
| a   | 7  | 20   | 14   | 24      | 72      | 50      |

7.4 z 值(见 GB/T 15483.1—1999,A.2.1.4,d)

7.4.1 总则

沿用 7.2 中符号,z 值定义为:

$$z = (x - \bar{X}) / \hat{\sigma} \dots\dots\dots (19)$$

其中  $\hat{\sigma}$  为能力评定标准差。

注: GB/T 15483.1 中用  $s$  表示  $z$  值定义中的标准差。当这个值是一些结果的标准差时,这个符号是适合的,而在其他情况下不适用(例如,当它是由精密度试验结果计算的标准差,或是如 6.4 所述的一般模型得出时)。

7.4.2 z 值的解释

当参加者报告结果的  $z$  值大于 3.0 或小于-3.0 时,则发出“行动信号”。同样的,当  $z$  值大于 2.0 或小于-2.0 时发出“警戒信号”。单个“行动信号”,或连续两轮计划中的“警戒信号”,应认为已出现异常,需进行调查。

若参加能力验证计划的实验室数量很多时(如超过 100 个实验室),如 7.9 中图 6 的正态概率图可补充解释  $z$  值。另一极端情况下,当实验室数很少时(如少于 10 个实验室),可能没有信号给出。这种情况下,应使用结合了多轮计划结果所得  $z$  值的图示法,将会提供比单轮计划结果更多的有关实验室性能的信息。

7.4.3 示例:抗体浓度的测定

表 7 中给出了由表 2 中稳健平均值和稳健标准差计算的  $z$  值,图 5 是其茎叶图。比较图 2 和图 5 可知  $z$  值和原始数据有相同的分布(不考虑修约的微小影响),但  $z$  值的中心为零点,与实验室偏倚的情况相同。

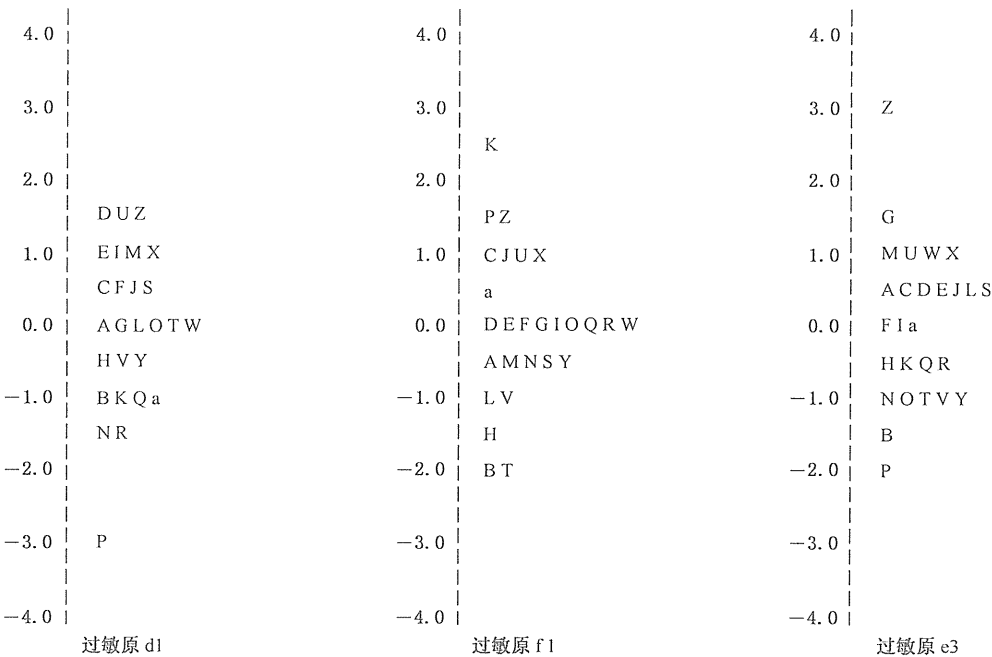


图 5 三种特定过敏原免疫球蛋白 E 抗体的浓度——z 值的茎叶图(由表 7 数据)



表 7 三种特定过敏原免疫球蛋白 E 抗体的浓度——z 值

| 实验室 | z 值    |        |       |
|-----|--------|--------|-------|
|     | d1     | f1     | e3    |
| A   | 0.09   | −0.28  | 0.54  |
| B   | −0.90  | −2.18W | −1.46 |
| C   | 0.29   | 0.80   | 0.64  |
| D   | 1.50   | −0.14  | 0.64  |
| E   | 0.78   | 0.16   | 0.39  |
| F   | 0.48   | −0.24  | 0.15  |
| G   | −0.21  | 0.10   | 1.27  |
| H   | −0.54  | −1.38  | −0.68 |
| I   | 1.04   | −0.18  | 0.10  |
| J   | 0.35   | 1.12   | 0.32  |
| K   | −0.96  | 2.54W  | −0.52 |
| L   | −0.08  | −0.88  | 0.28  |
| M   | 0.91   | −0.62  | 0.99  |
| N   | −1.33  | −0.66  | −0.76 |
| O   | −0.06  | −0.06  | −1.24 |
| P   | −2.91W | 1.38   | −1.98 |
| Q   | −0.87  | 0.00   | −0.44 |
| R   | −1.34  | 0.18   | −0.66 |
| S   | 0.25   | −0.50  | 0.41  |
| T   | −0.04  | −2.06W | −1.24 |
| U   | 1.73   | 1.12   | 1.00  |
| V   | −0.43  | −1.24  | −0.82 |
| W   | −0.17  | 0.20   | 0.80  |
| X   | 0.85   | 0.80   | 0.94  |
| Y   | −0.31  | −0.40  | −0.94 |
| Z   | 1.66   | 1.72   | 3.10A |
| a   | −0.84  | 0.66   | 0.23  |

注 1：表中 A=行动信号，W=警戒信号。

注 2：表中 z 值由表 2 中数据计算得到，其中将表 2 数据的稳健均值作为三个水平的指定值，并将表 2 的稳健标准差作为能力评定标准差。因而，本例中计算 z 值的公式为  $z = (x - \bar{X}) / \hat{\sigma} = (x - x^*) / s^*$ 。

7.5  $E_n$  值(见 GB/T 15483.1—1999, A.2.1.4, e)

性能统计量  $E_n$  值定义为:

$$E_n = \frac{x - X}{\sqrt{U_{\text{lab}}^2 + U_{\text{ref}}^2}} \quad \dots\dots\dots (20)$$

其中:

$X$  ——参照实验室确定的指定值;

$U_{\text{ref}}$  —— $X$  的扩展不确定度;

$U_{\text{lab}}$  ——参加者结果  $x$  的扩展不确定度。

相对于  $z$  值判定准则的临界值 2.0 和 3.0, 通常用于  $E_n$  值的临界值为 1.0。这是因为  $E_n$  值的计算是由扩展不确定度代替分母中的标准差得到。

注 1: 当参加者可能对其不确定度理解不足, 且不能以统一的形式做出报告时, 应谨慎使用  $E_n$  值。尽管如此, 将不确定度的信息与能力验证结果的解释相结合, 这对加强对这个问题的理解有很大帮助。

当用扩展因子 2.0 计算扩展不确定度时,  $E_n$  值的临界值 1.0 等价于  $z$  值的临界值 2.0。

注 2: 当按照 JJF 1059《测量不确定度评定与表示》评定不确定度时,  $E_n$  值表明了每个结果的扩展不确定度估计值的有效性。若  $|E_n| < 1$ , 则客观上表明不确定度估计值与 JJF 1059 中扩展不确定度的定义是一致的。

7.6  $z'$  值

## 7.6.1 总则

沿用 7.4 的符号,  $z'$  值定义如下:

$$z' = (x - X) / \sqrt{\hat{\sigma}^2 + u_x^2} \quad \dots\dots\dots (21)$$

其中  $u_x$  为指定值  $X$  的标准不确定度。

当指定值的计算未用到参加者的结果时, 可用式(21)来计算。因此, 当指定值由 5.2、5.3 和 5.4 中方法获得时, 或当应用了 5.5 中方法且专家实验室没有参加能力验证时, 可用式(21)。当使用 5.6 中方法时, 指定值与参加者的报告结果相关, 因而此时由式(21)定义的  $z'$  值是无效的。

7.6.2  $z'$  值的解释

$z'$  值与  $z$  值有相同的解释(见 7.4), 并有相同的临界值 2.0 和 3.0。

注: 7.1 和 7.2 中解释实验室偏倚和百分相对差的准则可作类似的修改, 即将  $\hat{\sigma}$  替换为  $\sqrt{\hat{\sigma}^2 + u_x^2}$ 。

7.6.3  $z'$  值的使用

比较 7.4 和 7.6 中  $z$  值和  $z'$  值的公式可知, 在一轮能力验证计划中  $z'$  值将小于相应的  $z$  值, 且相差一个常数因子

$$\hat{\sigma} / \sqrt{\hat{\sigma}^2 + u_x^2}$$

当 4.2 中指定值的不确定度的限制条件满足时, 这个因子将落在以下范围内:

$$0.96 \leq \hat{\sigma} / \sqrt{\hat{\sigma}^2 + u_x^2} \leq 1.00 \quad \dots\dots\dots (22)$$

因此, 这种情况下,  $z'$  值几乎与  $z$  值相同, 并且可以推断出指定值的不确定度是可以忽略的。

当 4.2 中条件不满足时,  $z'$  值与  $z$  值数量上的差异可能会导致: 一些  $z$  值超过了 2.0 或 3.0 的临界值因而发出“警戒信号”或“行动信号”, 然而相对应的  $z'$  值并没有超过这些临界值, 因而没有发出信号。

当选择使用  $z$  值或  $z'$  值时, 协调者应考虑以下几方面的问题。

- 指定值的不确定度是否满足 4.2 的准则? 若满足, 则使用  $z'$  值可能不会有任何帮助。
- 当该准则不满足时, 则建议使用  $z'$  值, 即使会增加复杂性并增加向使用者解释的难度。
- 当实验室结果显示出发警戒信号或行动信号时, 将会给实验室带来怎样严重的后果? 这些结果是否会导致该实验室失去为某些客户提供此类测试服务的资格?

## 7.7 $\zeta$ 值

### 7.7.1 总则

沿用 7.4 中符号,  $\zeta$  值定义为:

$$\zeta = (x - X) / \sqrt{u_x^2 + u_X^2} \quad \dots\dots\dots (23)$$

其中  $u_x$  是实验室自身对其结果  $x$  的标准不确定度的估计值,  $u_X$  是指定值  $X$  的标准不确定度。

当指定值的计算未用到参加者报告的结果时, 式(23)可用。因此, 当指定值由 5.2、5.3 和 5.4 中方法获得时, 并且当应用 5.5 中的方法且专家实验室没有参加能力验证时, 可用式(23)。当使用了 5.6 中方法时, 指定值与参加者的报告结果相关, 因此不能使用由式(23)定义的  $\zeta$  值。

注 1:  $\zeta$  值与  $E_n$  值的差别在于用标准不确定度代替了扩展不确定度。

注 2: 通常情况下, 能力验证计划中计算性能统计量时, 一般不会把参加实验室测量不确定度的信息包含在内。尽管如此, 参加实验室提供的信息仍然可能会在很多报告中给出。7.7 给出了协调者合并这些信息的指南。现在, GB/T 27025 要求说明测量不确定度的信息, 因而能力验证计划中的实验室应遵循该标准, 协调者需对怎样处理此类信息给出指导。

### 7.7.2 $\zeta$ 值的解释

当存在可以确认实验室对于自身结果的标准不确定度估计值的有效方法时,  $\zeta$  值可代替  $z$  值使用, 且应与  $z$  值用相同的方式进行解释(见 7.4), 使用相同的临界值 2.0 和 3.0。

当不存在这样的方法时,  $\zeta$  值只能与  $z$  值联合使用, 作为提升实验室性能的一种辅助。若实验室所得  $z$  值多次超出临界值 3.0, 则有必要逐步检查能力验证的程序, 并得到这个程序的不确定度来源一览表。不确定度来源一览表将识别能力验证程序中最大的不确定度来源, 因而实验室能找到改善结果的途径。若  $\zeta$  值仍然多次地超出临界值 3.0, 则表明不确定度来源一览表中未包含所有重要的不确定度来源(例如, 遗漏了某些重要因素)。

如果实验室存在一个较大偏倚, 并且它的不确定度区间  $x \pm U_x$  不包含指定值, 则它也将有较大的  $\zeta$  值或  $E_n$  值。

## 7.8 $E_x$ 值

$$E_x \text{ 值定义为: } E_{x-} = \frac{x - (X - U_X)}{U_x} \text{ 和 } E_{x+} = \frac{x - (X + U_X)}{U_x}。$$

其中  $X$  为与 5.4 或 5.6 的方法得到的指定值, 或者与 5.5 方法得到的参照值,  $U_X$  为  $X$  的扩展不确定度。  $x$  是实验室结果,  $U_x$  是  $x$  的扩展不确定度。  $U_x$  可以是 7.4 和 7.6 中的  $2\hat{\sigma}$  或  $3\hat{\sigma}$ , 及 7.5 中的  $U_{\text{lab}}$ 。

通常将  $E_x$  值与临界值 1.0 作比较:

- 当  $E_{x-}$  和  $E_{x+}$  均落在 -1.0 到 1.0 的范围内时, 实验室性能是令人满意的。
- 当  $E_{x-}$  和  $E_{x+}$  的其中一个值落在 -1.0 到 1.0 的范围之外时, 实验室性能是有问题的。
- 当  $E_{x-}$  和  $E_{x+}$  均小于 -1.0 或大于 1.0 时, 实验室性能是不可接受的。

## 7.9 已报告不确定度时进行数据分析的示例

### 7.9.1 总则

图 6 和图 7 给出了实验室结果及其不确定度的有用信息。出现最大偏倚的实验室有最大的  $z$  值(它将出现在图的两端)。

表 8 给出了 181 个实验室报告的水中铅含量的测试结果及其不确定度。报告数据以 mol/L 为单位, 数据均乘以  $10^{10}$  以方便处理。从表中可以看出这些结果既没有相同的不确定度也没有相同的偏倚。每个结果自身的偏倚记为  $D_i$ , 由方法偏倚和实验室偏倚之和得到; 此外, 每种方法有其自身的再现性标准差  $\sigma_R$ 。实验室可使用多种方法来确定测试结果, 并计算其不确定度。此处不确定度可认为是扩展不

确定度。表 8 中测试结果以升序排列,实验室也以此顺序编号。 $U=0$  说明无法报告不确定度。表中一些数据的有效性是可疑的,需要进一步调查,且不能在图中显示最极端的结果。特别地,负数值是不包含在图中的,尽管它们在计算中也会用到。这些数据给出了如 4.6 要求的报告负数结果的一个示例,尽管负的铅含量是不可能存在的。

### 7.9.2 指定值及其不确定度

使用如 5.6 中方法计算结果的稳健平均值作为指定值,应用附录 C 中算法 A。指定值为:

$$X = x^* = 605 \times 10^{-10} (\text{mol/L}) \quad \dots\dots\dots (24)$$

稳健标准差为:

$$s^* = 142 \times 10^{-10} (\text{mol/L}) \quad \dots\dots\dots (25)$$

依据 5.6,指定值的标准不确定度为:

$$u_x = 1.25 \times s^* / \sqrt{181} = 13 \times 10^{-10} (\text{mol/L}) \quad \dots\dots\dots (26)$$

### 7.9.3 能力评定标准差

能力评定标准差,如 6.6 所述,等于稳健标准差。

$$\hat{\sigma} = s^* = 142 \times 10^{-10} (\text{mol/L}) \quad \dots\dots\dots (27)$$

### 7.9.4 解释指定值不确定度的指南

根据 4.2 中准则,若

$$u_x \leq 0.3\hat{\sigma} \quad \dots\dots\dots (28)$$

指定值的不确定度将被忽略。当如示例中, $u_x = 1.25 \times s^* / \sqrt{p}$ , $\hat{\sigma} = s^*$  时,由计算可知, $p > 17$  时式(28)成立。当参与实验室数  $p = 181$  时,显然该准则易满足。因此这里无需考虑如 7.6 中的  $z'$  值。

### 7.9.5 利用正态概率图对大量实验室进行数据分析

图 6 显示了 181 个实验室结果关于实验室秩百分数(计算如 7.3 所述)的正态分布图。图中不包含小于 0 mol/L 或大于  $1\,600 \times 10^{-10}$  mol/L 的结果。

$z$  值可由  $z = (x - 605) / 142$  计算。当实验室获得的  $z$  值大于 3.0 或小于 -3.0 时,其  $z$  值已在图中对应的点附近标出。

均值为  $605 \times 10^{-10}$  mol/L 方差为  $142 \times 10^{-10}$  mol/L 的正态分布的累积分布函数在图中以虚直线标出。

最后,稳健算法中的截止值为:

$$x^* - 1.5s^* = 605 - 1.5 \times 142 = 392 \times 10^{-10} \text{ mol/L} \quad \dots\dots\dots (29)$$

和

$$x^* + 1.5s^* = 605 + 1.5 \times 142 = 818 \times 10^{-10} \text{ mol/L} \quad \dots\dots\dots (30)$$

在图 6 中可以看到偏离了虚线的数值都在这个范围之外。这表明所有的结果不是来源于同一个正态总体。偏离虚线的点来自方差更大的总体。

图中也可看到  $z$  值大于 3.0 或小于 -3.0 的实验室结果都偏离了虚线。也就支持了这些  $z$  值可导出“行动”信号这一结论。(大量结果的  $z$  值大于 3.0 或小于 -3.0 的情况也是可能发生的,但是,如图 6 所示所有的点都是靠近虚线的,这种情况下,由图形就不能得出  $z$  值产生行动信号的结论。)

图 7 显示了  $z$  值在  $\pm 3.0$  范围内的实验室结果,垂直线的长度显示实验室报告结果的不确定度。(24 号和 112 号实验室报告的结果的不确定度非常大,没有包含在图中。)图中水平实线代表指定值,两侧的虚线代表指定值的扩展不确定度。

若实验室对其结果的扩展不确定度进行了有效计算,则图 7 中几乎所有的垂直线将穿过指定值的扩展不确定度所围成的区域。但是,由图示可知很多实验室的结果均不能从两侧达到这个区域。这表明有很多实验室没有进行有效的不确定度计算;或者更有可能的是,它们的计算中没有包含所有重要的不确定度来源。

表 8 181 个实验室报告的水中铅含量( $\text{mol/L} \times 10^{-10}$ )测试结果及其扩展不确定度( $U$ )

| 实验室 | 实验结果     | $U$       | 实验室 | 实验结果 | $U$ | 实验室 | 实验结果 | $U$   | 实验室 | 实验结果              | $U$              |
|-----|----------|-----------|-----|------|-----|-----|------|-------|-----|-------------------|------------------|
| 1   | —960 000 | 0         | 51  | 545  | 43  | 101 | 618  | 224   | 151 | 740               | 20               |
| 2   | —12 100  | 0         | 52  | 545  | 123 | 102 | 618  | 170   | 152 | 748               | 3                |
| 3   | —4 800   | 0         | 53  | 550  | 8   | 103 | 620  | 25    | 153 | 767               | 113              |
| 4   | —3 860   | 0         | 54  | 550  | 55  | 104 | 620  | 40    | 154 | 772               | 213              |
| 5   | —1 500   | 0         | 55  | 550  | 5   | 105 | 621  | 6     | 155 | 800               | 60               |
| 6   | —1 010   | 0         | 56  | 555  | 79  | 106 | 622  | 9     | 156 | 800               | 150              |
| 7   | —1 000   | 0         | 57  | 556  | 30  | 107 | 622  | 6     | 157 | 821               | 203              |
| 8   | —1 000   | 0         | 58  | 557  | 28  | 108 | 623  | 18    | 158 | 830               | 10               |
| 9   | —965     | 0         | 59  | 557  | 28  | 109 | 625  | 15    | 159 | 857               | 27               |
| 10  | —438     | 0         | 60  | 559  | 26  | 110 | 626  | 5     | 160 | 874               | 200              |
| 11  | 160      | 20        | 61  | 560  | 7   | 111 | 627  | 0     | 161 | 898               | 59               |
| 12  | 180      | 20        | 62  | 560  | 60  | 112 | 627  | 1 010 | 162 | 900               | 100              |
| 13  | 203      | 0         | 63  | 569  | 116 | 113 | 627  | 15    | 163 | 920               | 140              |
| 14  | 256      | 13        | 64  | 570  | 86  | 114 | 628  | 3     | 164 | 950               | 110              |
| 15  | 319      | 0         | 65  | 571  | 16  | 115 | 629  | 26    | 165 | 965               | 0                |
| 16  | 335      | 18        | 66  | 572  | 40  | 116 | 630  | 40    | 166 | 968               | 0                |
| 17  | 340      | 180       | 67  | 574  | 35  | 117 | 630  | 580   | 167 | 990               | 0                |
| 18  | 400      | 20        | 68  | 578  | 0   | 118 | 632  | 50    | 168 | 990               | 80               |
| 19  | 404      | 36        | 69  | 579  | 52  | 119 | 637  | 96    | 169 | 1 010             | 0                |
| 20  | 407      | 0         | 70  | 579  | 35  | 120 | 630  | 83    | 170 | 1 250             | 140              |
| 21  | 410      | 97        | 71  | 579  | 8   | 121 | 640  | 130   | 171 | 1 320             | 410              |
| 22  | 444      | 58        | 72  | 579  | 10  | 122 | 640  | 77    | 172 | 1 450             | 460              |
| 23  | 450      | 20        | 73  | 579  | 17  | 123 | 642  | 20    | 173 | 1 640             | 241              |
| 24  | 450      | 3 400 000 | 74  | 579  | 87  | 124 | 647  | 63    | 174 | 1 900             | 46               |
| 25  | 463      | 19        | 75  | 580  | 150 | 125 | 647  | 0     | 175 | 2 413             | 20               |
| 26  | 470      | 10        | 76  | 582  | 122 | 126 | 650  | 160   | 176 | 2 460             | 0                |
| 27  | 470      | 30        | 77  | 589  | 57  | 127 | 650  | 30    | 177 | 2 900             | 900              |
| 28  | 474      | 0         | 78  | 589  | 10  | 128 | 650  | 80    | 178 | 10 000            | 0                |
| 29  | 480      | 100       | 79  | 590  | 0   | 129 | 650  | 48    | 179 | 386 000           | 31 000           |
| 30  | 482      | 122       | 80  | 590  | 0   | 130 | 650  | 30    | 180 | 670 000           | 60 000           |
| 31  | 483      | 241       | 81  | 590  | 45  | 131 | 653  | 5     | 181 | $630 \times 10^6$ | $60 \times 10^6$ |
| 32  | 490      | 60        | 82  | 590  | 60  | 132 | 658  | 27    |     |                   |                  |
| 33  | 492      | 25        | 83  | 590  | 0   | 133 | 660  | 20    |     |                   |                  |
| 34  | 492      | 1         | 84  | 591  | 112 | 134 | 660  | 120   |     |                   |                  |
| 35  | 493      | 24        | 85  | 591  | 9   | 135 | 660  | 34    |     |                   |                  |
| 36  | 493      | 5         | 86  | 594  | 4   | 136 | 663  | 32    |     |                   |                  |
| 37  | 495      | 0         | 87  | 594  | 119 | 137 | 675  | 280   |     |                   |                  |
| 38  | 500      | 70        | 88  | 597  | 9   | 138 | 675  | 0     |     |                   |                  |
| 39  | 500      | 10        | 89  | 600  | 20  | 139 | 680  | 8     |     |                   |                  |
| 40  | 500      | 10        | 90  | 600  | 300 | 140 | 680  | 50    |     |                   |                  |
| 41  | 501      | 75        | 91  | 603  | 60  | 141 | 680  | 70    |     |                   |                  |
| 42  | 504      | 0         | 92  | 603  | 24  | 142 | 685  | 0     |     |                   |                  |
| 43  | 510      | 130       | 93  | 603  | 13  | 143 | 700  | 0     |     |                   |                  |
| 44  | 510      | 110       | 94  | 604  | 18  | 144 | 700  | 100   |     |                   |                  |
| 45  | 512      | 6         | 95  | 608  | 30  | 145 | 700  | 110   |     |                   |                  |
| 46  | 526      | 26        | 96  | 608  | 8   | 146 | 700  | 300   |     |                   |                  |
| 47  | 530      | 9         | 97  | 609  | 9   | 147 | 708  | 44    |     |                   |                  |
| 48  | 530      | 40        | 98  | 610  | 61  | 148 | 709  | 48    |     |                   |                  |
| 49  | 530      | 60        | 99  | 613  | 22  | 149 | 710  | 100   |     |                   |                  |
| 50  | 545      | 30        | 100 | 618  | 7   | 150 | 729  | 41    |     |                   |                  |

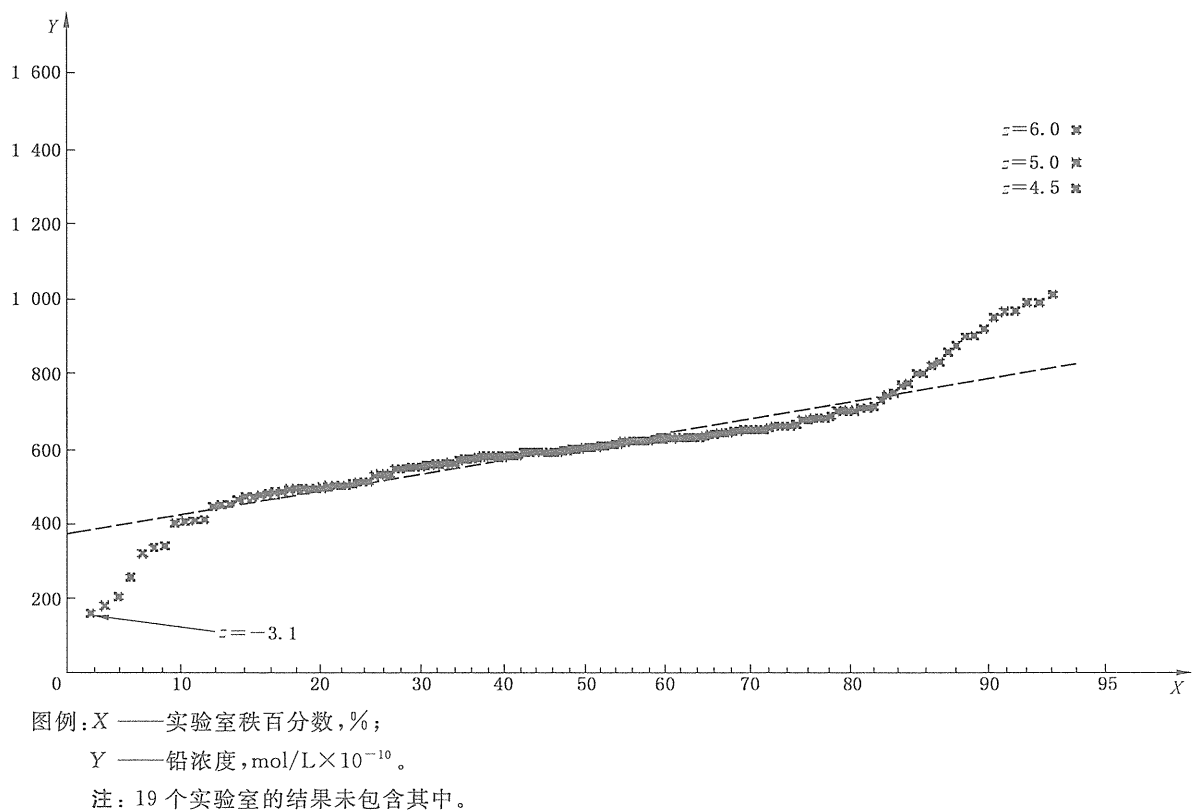


图 6 162 个实验室水中铅含量测定结果的正态概率图

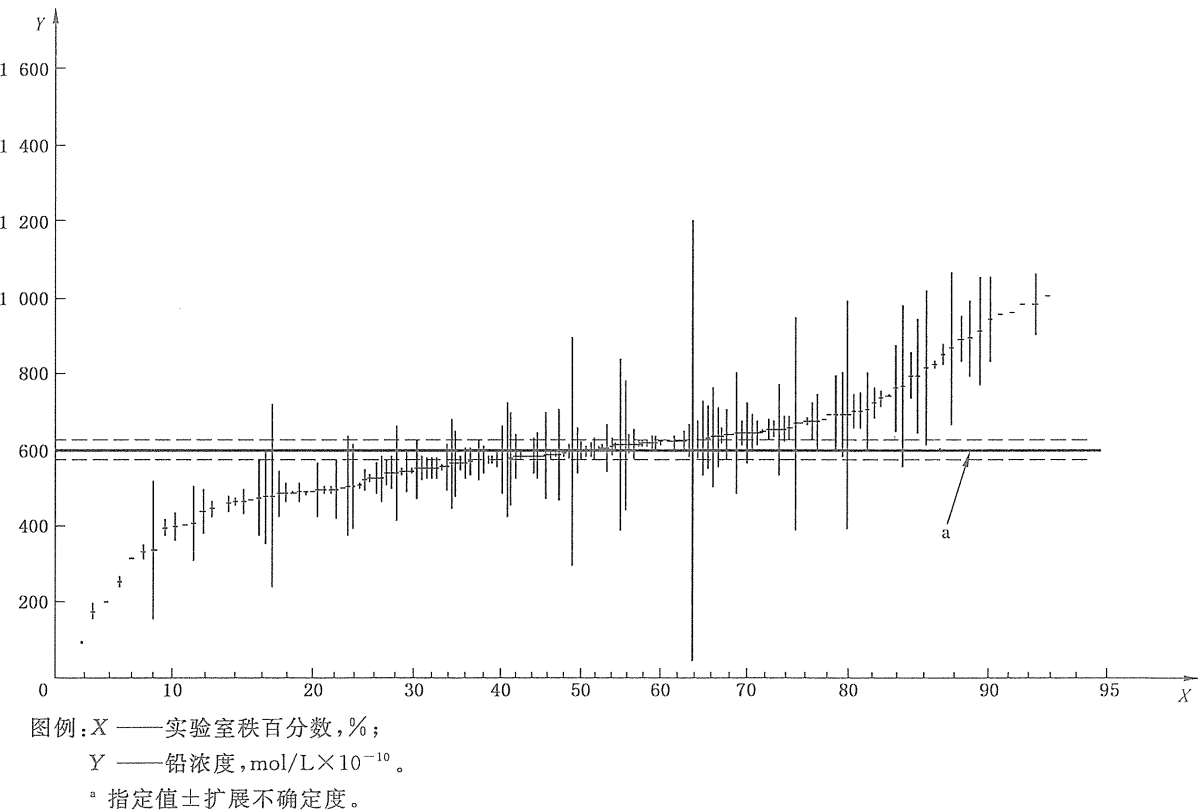


图 7 156 个实验室水中铅含量测定结果扩展不确定度的正态概率图

7.10 组合性能统计量值

在单轮能力验证计划中,通常会测得多个测试单元或被测量的结果。这种情况下,每个测试单元或被测量的结果应如 7.2 至 7.9 的方法分别进行解释,也就是说,应对每个测试单元或被测量的结果分别进行分析。当能力验证计划中需对两种或多种材料的特定水平进行测量,以确定其他方面性能(如,重复性或线性)时,协调者应为参加者提供统计设计和所用程序的完整说明。若计划中将两种相似材料组合作为尧敦对(Youden Pair),应使用 8.5 中的特殊方法。

当对多个测试单元或同一测试单元的多个被测量进行测试并得到多个结果时,推荐使用第 8 章的某些图示法。这种方法将结合性能统计量值而不隐藏单个性能统计量值中的较高值,并会显示出在统计量值表中并不明显的一些实验室性能的额外信息,比如,不同被测量间的相关性。

当出现一定数量的行动信号和警戒信号时,应识别产生这些信号的实验室,并向这些实验室提供一份详细报告,该报告应包含使用 7.2 至 7.9 中方法所得结果的细节。对于没有任何信号的实验室,仅提供一份简要报告。

注:当对多个测试单元有相同的被测量时,不应使用  $z$  值之和或其平均值、差的绝对值或差的平方的和或平均值等组合性能统计量值;或者存在多个被测量的结果时,也不应使用  $z$  值绝对值的平均值、或评价限的平均绝对差值等组合性能统计量值。当一个测试单元存在高的性能统计量值,但其他性能统计量值很低,或另一个性能统计量值也很高但有相反的符号时,此时  $z$  值的平均值(或和)会降低。当一个测试单元存在一个高性能统计量值,而其他性能统计量值很低时,此时差的绝对值和差的平方的平均值(或和)也会降低。 $z$  值绝对值的平均值和关于评价限的相对差的绝对值的平均值也会有相同的问题。

8 一轮能力验证计划中多个被测量的组合性能统计量值的图示法

8.1 应用

协调者应按照 8.2 和 8.3 所述,使用每轮能力验证计划所获的性能统计量值来画图。这些图中使用  $z$  值,其优点在于这些点可在标准坐标轴中描出,因而简化了它们的表示和解释。图的绘制应方便参加者,使每个参加者能在图中看到自己的结果与其他参加者所得结果的关系。可用字母编号或数字编号代表参加者,以便每个参加者能辨别自己的结果但不能获得其他参加者的结果。图示还应方便协调者,使他们能判断整个能力验证计划的有效性,并确定是否需要重新审查能力评估的标准。

8.2 性能统计量值的直方图

8.2.1 总则

要绘制  $z$  值的直方图,需收集一轮能力验证计划中一种特性测量值的  $z$  值,如图 8 所示。直方图组距设为 0.3~0.5,以便图示清晰。在  $\pm 2.0$  和  $\pm 3.0$  处的直线,代表能力评估的标准。直方图的范围定为  $\pm 6.0$ 。若所得结果没有超出这个范围,或紧密集集中在中间部位,分析员可考虑其他合适的范围。

当需要绘制实验室偏倚或百分相对差的直方图时,与  $z$  值直方图等价的直方图组距、警戒限和行动限如表 9 所示。在这些情况中,直接由规定值或经验预期值来确定行动限比由能力验证的标准差来计算行动限可能要简单一些。

表 9 性能统计量值的警戒限和行动限

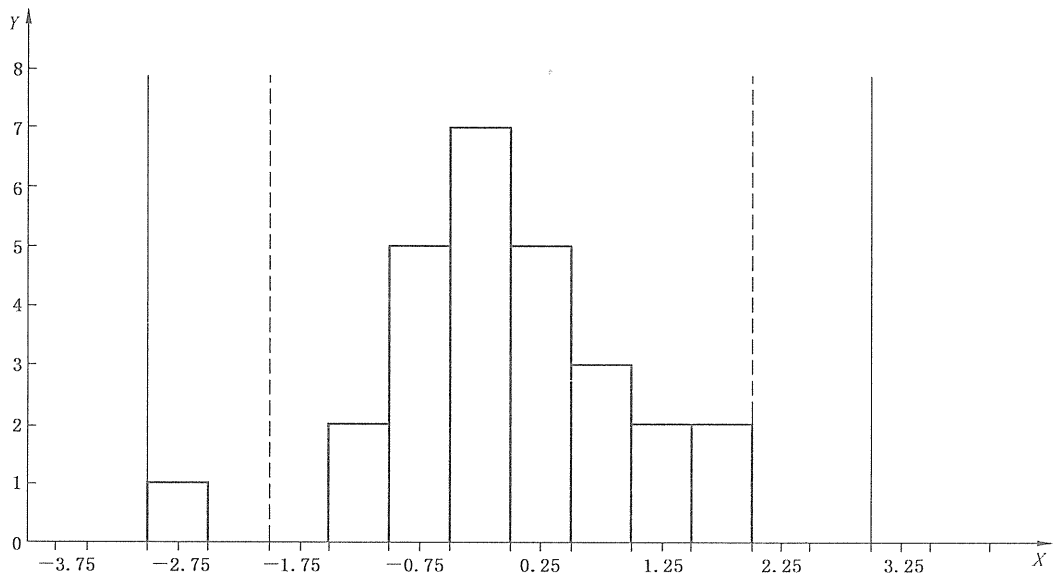
| 性能统计量 | 直方图组距                                    | 警戒限                     | 行动限                     | 直方图范围                   |
|-------|------------------------------------------|-------------------------|-------------------------|-------------------------|
| 实验室偏倚 | $0.3\hat{\sigma} \sim 0.5\hat{\sigma}$   | $\pm 2.0\hat{\sigma}$   | $\pm 3.0\hat{\sigma}$   | $\pm 6.0\hat{\sigma}$   |
| 百分相对差 | $30\hat{\sigma}/X \sim 50\hat{\sigma}/X$ | $\pm 200\hat{\sigma}/X$ | $\pm 300\hat{\sigma}/X$ | $\pm 600\hat{\sigma}/X$ |
| $z$ 值 | 0.3~0.5                                  | $\pm 2.0$               | $\pm 3.0$               | $\pm 6.0$               |

当被测特性的数量很少,或需测量一些不类似的特性时,直方图是一种合适的图示法。每个参加者能确定自己所得分值的位置,评估其实验能力并且判断是否需要检查其测量方法。获得高  $z$  值的参加者能够通过直方图与其他参加者进行比较,来确定其异常程度。

协调者能够利用直方图确定有多少参加者不能满足能力评估标准。若直方图的尾部超出了一3.0~+3.0的比例较多,则表明原因可能在于使用的测量方法,而不在于参加者,因此需要改进测量方法,或放宽能力评估标准(即,增大  $\hat{\sigma}$ )。若直方图均在-2.0~+2.0 的范围以内,或者仅有一个或两个  $z$  值在这个范围之外,则建议加严能力评估标准(即,减小  $\hat{\sigma}$ )。

8.2.2 示例:抗体浓度的测定

图 8 是 d1 的  $z$  值的直方图。



图例: X —— 过敏原 d1 的  $z$  值;  
Y —— 实验室数。

图 8 一轮能力验证计划中  $z$  值的直方图(由表 7 中过敏原 d1 的数据)

8.3 标准化实验室偏倚的条形图

8.3.1 总则

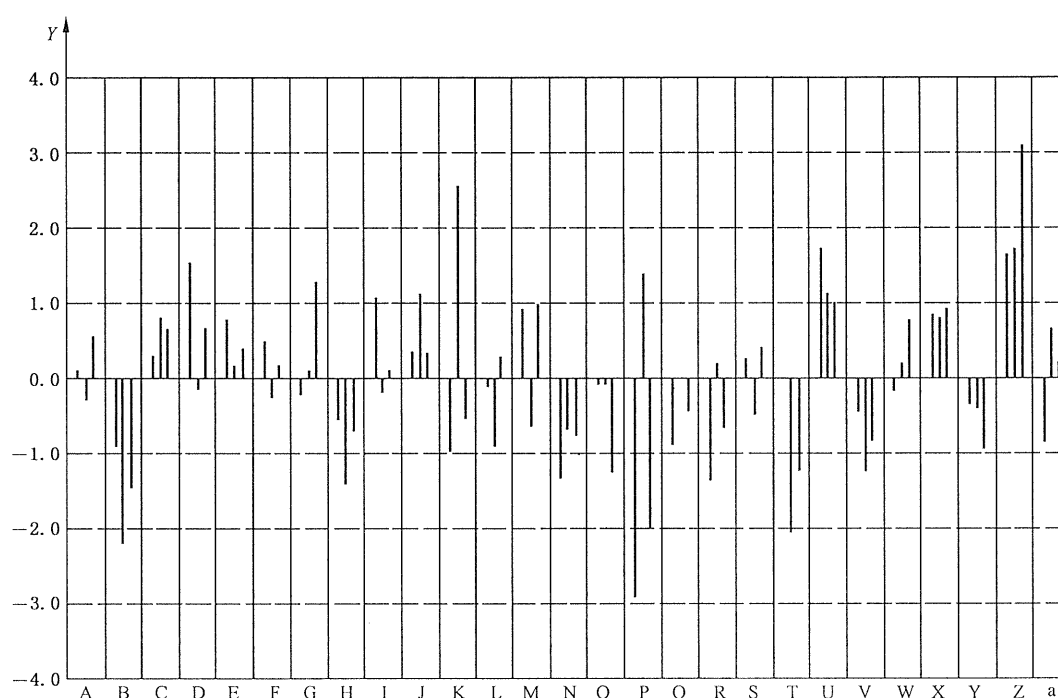
要绘制条形图,需收集  $z$  值绘入图中,如图 9 所示,每位参加者的 3 个  $z$  值列为一组。(  $z$  值与 GB/T 6379.2 中提到的  $h$  统计量是相同的,此处的条形图与 GB/T 6379.2 中  $h$  统计量图相同。)

条形图适合于将一些相似特性的  $z$  值表示在同一图中,它可以显示出某个参加者的所有  $z$  值是否存在一些共同特征。例如,若某个参加者得到一些较高的  $z$  值则表明其性能较差。

8.3.2 示例:抗体浓度的测定

表 7 数据的  $z$  值条形图如图 9 所示。由图可知,实验室 B 和 Z 应检查在 3 个水平上产生大致相同的实验室偏倚的原因,而由实验室 K 和 P 的结果可知其  $z$  值的符号依赖于测试抗体的种类。





图例:Y—— $z$  值。

图 9 一轮能力验证计划中  $z$  值(4.0~−4.0)的条形图,  
其中参加者测定了三种特定过敏原免疫球蛋白 E 抗体的浓度(由表 7 数据)

#### 8.4 标准化重复性测量值的条形图

当一轮能力验证计划中对被测量进行了重复测试,其结果可用于计算  $k$  统计量并绘图,如 GB/T 6379.2 所述。

#### 8.5 尧敦图

##### 8.5.1 总则

当两种相似物质的样品已在一轮能力验证计划中进行了测试,尧敦图提供了研究这些结果的非常有效的图示法。图中  $X$  轴为一种物质的  $z$  值, $Y$  轴为另一种物质的  $z$  值。根据 8.5.2 计算置信椭圆,用于辅助解释该图。对于原始数据,实验室偏倚或偏倚百分比的尧敦图可根据本节中注 1,由  $z$  值得出。

尧敦图的解释如下。

- 检查图中远离其他数据的点。若实验室没有正确遵循测试方法,其结果将存在较大偏倚,会出现远离椭圆主轴的点。当实验室在其结果水平上经常存在较大变异时,也会出现远离主轴的点。远离主轴的点代表重复性较差的参加者。
- 观察图中是否有证据表明两种物质的测试结果存在某种关系。若存在,则表明对许多实验室来说,普遍存在引起实验室间变异的原因,并且表明测量方法规范的不够详细。检查测试方法一般可以改善方法的再现性。8.5.3 中给出的秩相关检验可用于检验两种物质间的相关性是否统计显著。这里推荐使用秩相关系数,因为传统的相关系数对数据的非正态性更敏感。

8.5.2 置信椭圆(基于杰克逊(Jackson)方法<sup>[2]</sup>)

## 8.5.2.1 总则

用 A 和 B 表示两种物质, A 的测试结果为:

$x_{A,1}, x_{A,2}, \dots, x_{A,p}$ ,

B 的测试结果为:

$x_{B,1}, x_{B,2}, \dots, x_{B,p}$ ,

其中  $p$  为实验室数。

计算这两组数据的均值、标准差和相关系数:

$\bar{x}_{A..}, \bar{x}_{B..}, s_A, s_B, \hat{\rho}$ 。

计算两种物质的  $z$  值:

$$z_{A,i} = (x_{A,i} - \bar{x}_{A..})/s_A, \quad i = 1, 2, \dots, p \quad (31)$$

$$z_{B,i} = (x_{B,i} - \bar{x}_{B..})/s_B, \quad i = 1, 2, \dots, p \quad (32)$$

计算两种物质的组合性能统计量值:

$$z_{A,B,i} = \sqrt{z_{A,i}^2 - 2\hat{\rho}z_{A,i}z_{B,i} + z_{B,i}^2} \quad (33)$$

定义标准化变量:

$$z_A = (x_A - \bar{x}_{A..})/s_A \quad (34)$$

$$z_B = (x_B - \bar{x}_{B..})/s_B \quad (35)$$

置信椭圆方程可用标准化变量和霍特林(Hotelling)  $T^2$  来表示:

$$z_A^2 - 2\hat{\rho}z_Az_B + z_B^2 = (1 - \hat{\rho}^2)T^2 \quad (36)$$

其中

$$T^2 = 2\{(p-1)/(p-2)\}F_{(1-\alpha)}(2, p-1) \quad (37)$$

此处,  $F_{(1-\alpha)}(2, p-1)$  表示自由度为 2 和  $(p-1)$  的 F-分布的  $(1-\alpha)$  分位数。置信度为  $(1-\alpha)$  的置信椭圆绘制在横坐标为  $z_A$  纵坐标为  $z_B$  的坐标系中, 可按下式画出:

$$z_B = \hat{\rho}z_A \pm \sqrt{(1 - \hat{\rho}^2)(T^2 - z_A^2)}, \quad -T \leq z_A \leq T \quad (38)$$

注 1: 若要在以原始测量单位为坐标轴的图中绘制置信椭圆, 可用下式将上述的点转换为原始单位:

$$x_A = \bar{x}_{A..} + s_A \times z_A$$

$$x_B = \bar{x}_{B..} + s_B \times z_B$$

若要在以实验室偏倚  $D_A$  和  $D_B$  为坐标轴的图中绘制置信椭圆, 可用下式将点转换为实验室偏倚:

$$D_A = s_A \times z_A$$

$$D_B = s_B \times z_B$$

若要在以百分相对差  $D_{A\%}$  和  $D_{B\%}$  为坐标轴的图中绘制置信椭圆, 可用下式将点转化为百分相对差:

$$D_{A\%} = 100s_A \times z_A/x_A$$

$$D_{B\%} = 100s_B \times z_B/x_B$$

组合  $z$  值可以辅助解释尧敦图。在置信椭圆的计算中, 最高的组合  $z$  值对应于最高的显著性水平  $\alpha$ 。因此组合  $z$  值可以确定尧敦图最极端的点。有时, 有必要剔除一个或多个离群点并重新计算椭圆: 此时可用组合  $z$  值确定可剔除的点。

注 2: 这里有必要利用稳健方法计算椭圆, 但目前还没有该方法的具体步骤。注意到  $(z_{A,B,i})^2/(1-\hat{\rho}^2)$  近似服从自由度为 2 的卡方分布, 可由此计算截止值, 但修正因子可能需由统计模拟得到。

## 8.5.2.2 示例: 抗体浓度的测定

表 10 给出了检测两种相似样本中抗体浓度的数据, 以及导出置信椭圆的相关量。实验室数  $p=29$ ,

显著性水平  $\alpha=5\%$ ,  $F_{(1-\alpha)}(2, p-1)=3.34$ , 因此  $T=2.632$ 。利用标准化变量, 置信水平为 95% 的置信椭圆方程为:

$$z_A^2 - 1.412z_A z_B + z_B^2 = 3.48 \quad \dots\dots\dots (39)$$

图 10 给出了置信水平为 95%、99% 和 99.9% 的置信椭圆, 图中的点代表  $z$  值。组合性能统计量值在表 10 中给出。

图 10 显示两个实验室(编号为 5 和 23)的  $z$  值在图中右上象限, 其  $z$  值组合性能统计量值为 1.641 和 2.099。26 号实验室对 B 物质有一个高  $z$  值, 其组合  $z$  值为 2.059。8 号实验室出现了除 5、23 和 26 号实验室之外的数据中最高的组合  $z$  值(值为 1.501)。

23 和 26 号实验室的点落在置信水平为 95% 和 99% 的椭圆之间, 因而这两个结果可作为警戒信号处理, 并在下轮计划中检查这两个实验室的结果落在何处。

表 10 两种相似过敏原抗体浓度的数据和计算

| 实验室<br><i>i</i> | 数据                      |                         | <i>z</i> 值              |                         | 组合 <i>z</i> 值<br><i>z</i> <sub>A,B,i</sub> |
|-----------------|-------------------------|-------------------------|-------------------------|-------------------------|--------------------------------------------|
|                 | 过敏原<br>A                | 过敏原<br>B                | 过敏原<br>A                | 过敏原<br>B                |                                            |
|                 | <i>x</i> <sub>A,i</sub> | <i>x</i> <sub>B,i</sub> | <i>z</i> <sub>A,i</sub> | <i>z</i> <sub>B,i</sub> |                                            |
| 1               | 12.95                   | 9.15                    | 0.427                   | 0.515                   | 0.370                                      |
| 2               | 6.47                    | 6.42                    | -1.540                  | -0.428                  | 1.275                                      |
| 3               | 11.40                   | 6.60                    | -0.043                  | -0.366                  | 0.336                                      |
| 4               | 8.32                    | 4.93                    | -0.978                  | -0.942                  | 0.737                                      |
| 5               | 18.88                   | 13.52                   | 2.228                   | 2.023                   | 1.641                                      |
| 6               | 15.14                   | 8.22                    | 1.092                   | 0.194                   | 0.965                                      |
| 7               | 10.12                   | 7.26                    | -0.432                  | -0.138                  | 0.349                                      |
| 8               | 17.94                   | 9.89                    | 1.942                   | 0.770                   | 1.501                                      |
| 9               | 11.68                   | 4.17                    | 0.042                   | -1.204                  | 1.234                                      |
| 10              | 12.44                   | 7.39                    | 0.272                   | -0.093                  | 0.344                                      |
| 11              | 6.93                    | 7.78                    | -1.400                  | 0.042                   | 1.430                                      |
| 12              | 9.57                    | 5.80                    | -0.599                  | -0.642                  | 0.477                                      |
| 13              | 11.73                   | 5.77                    | 0.057                   | -0.652                  | 0.693                                      |
| 14              | 12.29                   | 6.97                    | 0.227                   | -0.238                  | 0.429                                      |
| 15              | 10.95                   | 6.23                    | -0.180                  | -0.493                  | 0.388                                      |
| 16              | 10.95                   | 5.90                    | -0.180                  | -0.607                  | 0.497                                      |
| 17              | 11.17                   | 7.74                    | -0.113                  | 0.028                   | 0.134                                      |
| 18              | 11.20                   | 8.63                    | -0.104                  | 0.335                   | 0.415                                      |
| 19              | 7.64                    | 3.74                    | -1.185                  | -1.353                  | 0.986                                      |
| 20              | 12.17                   | 7.33                    | 0.190                   | -0.114                  | 0.282                                      |
| 21              | 10.71                   | 5.70                    | -0.253                  | -0.676                  | 0.529                                      |
| 22              | 7.84                    | 6.07                    | -1.124                  | -0.549                  | 0.833                                      |
| 23              | 20.47                   | 15.66                   | 2.710                   | 2.762                   | 2.099                                      |
| 24              | 12.60                   | 11.76                   | 0.321                   | 1.415                   | 1.210                                      |

表 10 (续)

| 实验室<br><i>i</i> | 数据                             |                                | <i>z</i> 值                     |                                | 组合 <i>z</i> 值<br><i>z</i> <sub>A,B,<i>i</i></sub> |
|-----------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|---------------------------------------------------|
|                 | 过敏原<br>A                       | 过敏原<br>B                       | 过敏原<br>A                       | 过敏原<br>B                       |                                                   |
|                 | <i>x</i> <sub>A,<i>i</i></sub> | <i>x</i> <sub>B,<i>i</i></sub> | <i>z</i> <sub>A,<i>i</i></sub> | <i>z</i> <sub>B,<i>i</i></sub> |                                                   |
| 25              | 11.37                          | 4.91                           | −0.052                         | −0.949                         | 0.913                                             |
| 26              | 11.36                          | 13.51                          | −0.055                         | 2.019                          | 2.059                                             |
| 27              | 10.75                          | 5.48                           | −0.241                         | −0.752                         | 0.607                                             |
| 28              | 12.21                          | 9.77                           | 0.203                          | 0.729                          | 0.603                                             |
| 29              | 7.49                           | 5.82                           | −1.230                         | −0.635                         | 0.902                                             |
| 均值              | 11.54                          | 7.66                           | 0.00                           | 0.00                           |                                                   |
| 标准差             | 3.29                           | 2.90                           | 1.00                           | 1.00                           |                                                   |
| 相关系数            | 0.706                          |                                | 0.706                          |                                |                                                   |

注 1: 表中数据为每升(L)样品中千单位(kU)抗体数,其中单位(U)由国际标准物质的浓度定义。  
注 2: 表中 *z* 值的计算使用了未圆整的平均值和标准差数据,而不是上表底部给出了数据。

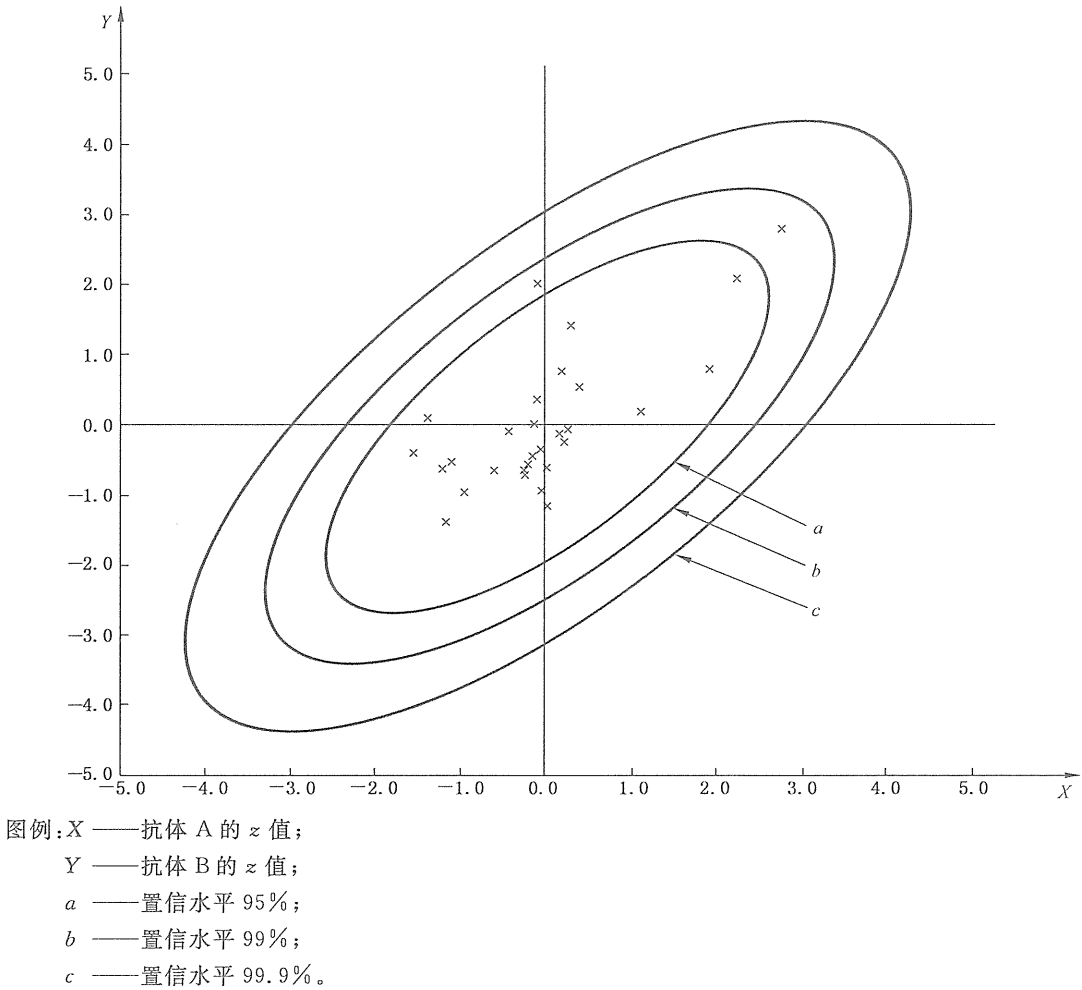


图 10 表 10 中数据的 *z* 值尧敦图

8.5.3 秩相关检验

8.5.3.1 总则

利用斯皮尔曼(Spearman)相关系数进行秩相关检验。将 A 物质的测试结果替换为其相应的秩(例如替换最小值为 1,第二小的为 2,以此类推,直到将最大值替换为  $p$ )。同样地处理 B 物质测试结果。如果数据中有结(两个或两个以上相同的观测结果),则这些相同结果的秩为它们的平均值。计算两组秩序列的相关系数,并与表 11 中相应的临界值比较。若计算所得相关系数大于临界值,则两个变量是显著相关的。秩相关系数可简单计算如下。记物质 A 的实验室测试结果的秩为:

$k_{A,1}, k_{A,2}, \dots, k_{A,p}$

物质 B 的实验室测试结果的秩为:

$k_{B,1}, k_{B,2}, \dots, k_{B,p}$

则秩相关系数可由下式计算:

$$\rho_k = 1 - 6 \sum (k_{A,i} - k_{B,i})^2 / \{p(p^2 - 1)\} \dots\dots\dots (40)$$

其中求和号是对  $p$  个实验室结果求和。

表 11 秩相关系数的临界值

| 数据组数 | 临界值      |          |
|------|----------|----------|
|      | 显著性水平 5% | 显著性水平 1% |
| 8    | 0.738    | 0.881    |
| 9    | 0.683    | 0.833    |
| 10   | 0.648    | 0.794    |
| 11   | 0.623    | 0.818    |
| 12   | 0.591    | 0.780    |
| 13   | 0.566    | 0.745    |
| 14   | 0.545    | 0.716    |
| 15   | 0.525    | 0.689    |
| 16   | 0.507    | 0.666    |
| 17   | 0.490    | 0.645    |
| 18   | 0.476    | 0.625    |
| 19   | 0.462    | 0.608    |
| 20   | 0.450    | 0.591    |
| 21   | 0.438    | 0.576    |
| 22   | 0.428    | 0.562    |
| 23   | 0.418    | 0.549    |
| 24   | 0.409    | 0.537    |
| 25   | 0.400    | 0.526    |
| 26   | 0.392    | 0.515    |
| 27   | 0.385    | 0.505    |
| 28   | 0.377    | 0.496    |
| 29   | 0.370    | 0.487    |
| 30   | 0.364    | 0.478    |

## 8.5.3.2 示例:抗体浓度的测定

表 12 给出了一个计算秩相关系数的示例。所得秩相关系数为 0.605,大于表 11 中数据组数为 29 显著性水平为 1%对应的临界值 0.487,表明图 10 显示的相关性是统计显著的。

表 12 表 10 数据的秩相关系数的计算

| 实验室<br>$i$                                     | 数据                    |                       | 秩                     |                       | 差值<br>$ k_{A,i} - k_{B,i} $ |
|------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------------|
|                                                | 过敏原<br>A<br>$z_{A,i}$ | 过敏原<br>B<br>$z_{B,i}$ | 过敏原<br>A<br>$k_{A,i}$ | 过敏原<br>B<br>$k_{B,i}$ |                             |
| 1                                              | 12.95                 | 9.15                  | 25                    | 23                    | 2                           |
| 2                                              | 6.47                  | 6.42                  | 1                     | 13                    | 12                          |
| 3                                              | 11.40                 | 6.60                  | 17                    | 14                    | 3                           |
| 4                                              | 8.32                  | 4.93                  | 6                     | 4                     | 2                           |
| 5                                              | 18.88                 | 13.52                 | 28                    | 28                    | 0                           |
| 6                                              | 15.14                 | 8.22                  | 26                    | 21                    | 5                           |
| 7                                              | 10.12                 | 7.26                  | 8                     | 16                    | 8                           |
| 8                                              | 17.94                 | 9.89                  | 27                    | 25                    | 2                           |
| 9                                              | 11.68                 | 4.17                  | 18                    | 2                     | 16                          |
| 10                                             | 12.44                 | 7.39                  | 23                    | 18                    | 5                           |
| 11                                             | 6.93                  | 7.78                  | 2                     | 20                    | 18                          |
| 12                                             | 9.57                  | 5.80                  | 7                     | 8                     | 1                           |
| 13                                             | 11.73                 | 5.77                  | 19                    | 7                     | 12                          |
| 14                                             | 12.29                 | 6.97                  | 22                    | 15                    | 7                           |
| 15                                             | 10.95                 | 6.23                  | 11.5                  | 12                    | 0.5                         |
| 16                                             | 10.95                 | 5.90                  | 11.5                  | 10                    | 1.5                         |
| 17                                             | 11.17                 | 7.74                  | 13                    | 19                    | 6                           |
| 18                                             | 11.20                 | 8.63                  | 14                    | 22                    | 8                           |
| 19                                             | 7.64                  | 3.74                  | 4                     | 1                     | 3                           |
| 20                                             | 12.17                 | 7.33                  | 20                    | 17                    | 3                           |
| 21                                             | 10.71                 | 5.70                  | 9                     | 6                     | 3                           |
| 22                                             | 7.84                  | 6.07                  | 5                     | 11                    | 6                           |
| 23                                             | 20.47                 | 15.66                 | 29                    | 29                    | 0                           |
| 24                                             | 12.60                 | 11.76                 | 24                    | 26                    | 2                           |
| 25                                             | 11.37                 | 4.91                  | 16                    | 3                     | 13                          |
| 26                                             | 11.36                 | 13.51                 | 15                    | 27                    | 12                          |
| 27                                             | 10.75                 | 5.48                  | 10                    | 5                     | 5                           |
| 28                                             | 12.21                 | 9.77                  | 21                    | 24                    | 3                           |
| 29                                             | 7.49                  | 5.82                  | 3                     | 9                     | 6                           |
| 差值平方和                                          |                       |                       |                       |                       | 1 605.50                    |
| $p(p^2 - 1)$                                   |                       |                       |                       |                       | 24 360                      |
| 秩相关系数                                          |                       |                       |                       |                       | 0.605                       |
| 1%显著性水平下的临界值                                   |                       |                       |                       |                       | 0.487                       |
| 注:表中数据为每升(L)样品中千单位(kU)抗体数,其中单位(U)由国际标准物质的浓度定义。 |                       |                       |                       |                       |                             |

8.6 重复性标准差图

8.6.1 总则

在一轮能力验证计划中参加者获得了  $n$  个重复测试结果时,利用这些结果的图示可以识别均值和标准差异异常的实验室。该图是对每个实验室的实验室内标准差  $s_i$  和实验室均值  $x_i$  分别描点得到。令

$\bar{X} = x^*$ , 其中  $x^*$  为用算法 A 计算的  $x_1, x_2, \dots, x_p$  的稳健平均值,

$\bar{S} = s^*$ , 其中  $s^*$  为用算法 S 计算的  $s_1, s_2, \dots, s_p$  的稳健联合值,

并假定数据服从正态分布。在原假设“各实验室的均值与实验室内标准差没有差异”成立时,统计量

$$\left[ \sqrt{n} \frac{x_i - \bar{X}}{\bar{S}} \right]^2 + \left[ \sqrt{2(n-1)} \ln \left( \frac{s_i}{\bar{S}} \right) \right]^2$$

近似服从自由度为 2 的  $\chi^2$  分布。因此,在以标准差  $s$  为纵轴、平均值  $x$  为横轴的坐标系中显著性水平为 1% 时的临界区域由下式确定:

$$s = \bar{S} \exp \left\{ \pm \frac{1}{\sqrt{2(n-1)}} \sqrt{\chi_{2;0.99}^2 - \left( \sqrt{n} \frac{x - \bar{X}}{\bar{S}} \right)^2} \right\}, \bar{X} - \bar{S} \sqrt{\frac{\chi_{2;0.99}^2}{n}} \leq x \leq \bar{X} + \bar{S} \sqrt{\frac{\chi_{2;0.99}^2}{n}} \quad \dots\dots\dots (41)$$

注:圆方法(circle technique)由 Van Nuland [6] 提出。该方法利用了标准差分布的简单正态近似,但该方法可能得出包含负标准差的临界区域。此处给出的方法是利用标准差的近似分布,从而避免了这种问题的发生,但此时临界区域不再是圆形了。另外,中心点使用稳健值代替了原方法中的简单平均值。

8.6.2 示例:抗体浓度的测定

表 13 为血浆样本中特定抗体浓度的测试结果。每个实验室在重复性条件下做四次重复测试。利用以上给出的公式可得图 11。该图表明一部分实验室收到了行动或警戒信号。

表 13 血浆样本中特定抗体的浓度  
(每个实验室对一份样本进行 4 次重复测定)

| 实验室 | 平均值<br>kU/L | 标准差<br>kU/L |
|-----|-------------|-------------|
| 1   | 2.15        | 0.13        |
| 2   | 1.85        | 0.21        |
| 3   | 1.80        | 0.08        |
| 4   | 1.80        | 0.24        |
| 5   | 1.90        | 0.36        |
| 6   | 1.90        | 0.32        |
| 7   | 1.90        | 0.14        |
| 8   | 2.05        | 0.26        |
| 9   | 2.35        | 0.39        |
| 10  | 2.03        | 0.53        |
| 11  | 2.08        | 0.25        |
| 12  | 1.25        | 0.24        |
| 13  | 1.13        | 0.72        |
| 14  | 1.00        | 0.26        |
| 15  | 1.08        | 0.17        |
| 16  | 1.20        | 0.32        |
| 17  | 1.35        | 0.40        |

表 13 (续)

| 室 验 室                                          | 平均值<br>kU/L | 标准差<br>kU/L |
|------------------------------------------------|-------------|-------------|
| 18                                             | 1.23        | 0.36        |
| 19                                             | 1.23        | 0.33        |
| 20                                             | 0.90        | 0.43        |
| 21                                             | 1.48        | 0.40        |
| 22                                             | 1.20        | 0.55        |
| 23                                             | 1.73        | 0.39        |
| 24                                             | 1.43        | 0.30        |
| 25                                             | 1.28        | 0.22        |
| 稳健均值                                           | 1.57        |             |
| 稳健标准差                                          |             | 0.34        |
| 注：表中数据为每升(L)样品中千单位(kU)抗体数，其中单位(U)由国际标准物质的浓度定义。 |             |             |

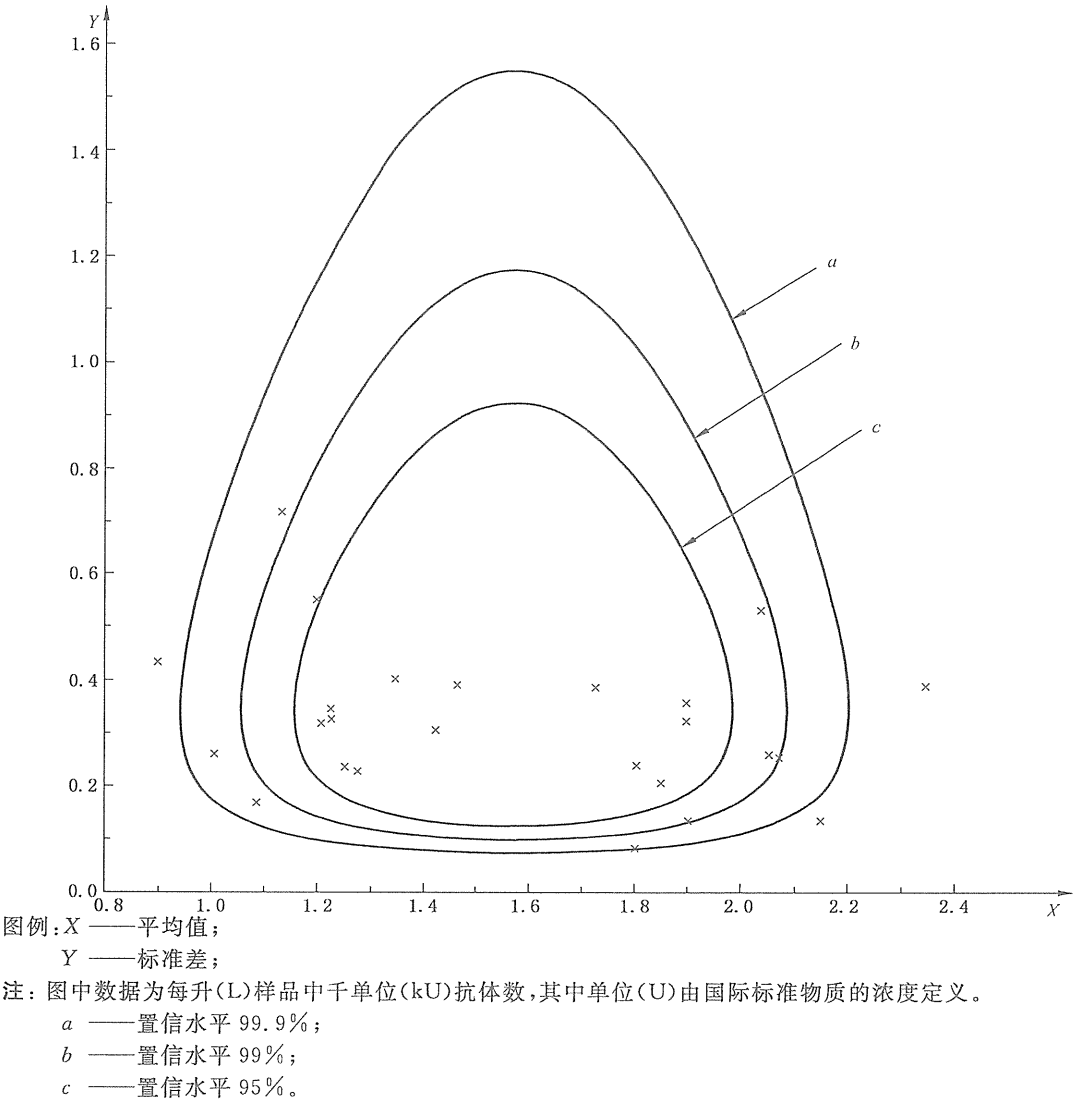


图 11 25 个实验室标准差和平均值图(由表 13 数据)



8.7 分割样品(见 GB/T 15483.1—1999,A.3.1.2)

8.7.1 总则

当有必要对两个实验室进行详细比对时,可使用分割样品。例如,当一个实验室由供应方管理,而另一个实验室由使用方管理,且这两个机构希望实验室结果保证一致。实验室将得到多种水平待测特性的样品,将每个样品分割为两份,并且每个实验室对每份分割样品进行若干次(至少两次)重复测定。

有时,进行比对的实验室会多于两个,这种情况下,可将一个实验室看作参照实验室,其他实验室应利用此处的技术与其进行比对。

由分割样品实验所得数据,可得到反映两个实验室重复测量的变异和每份样品平均结果差值的图示。依据对图示的推断做进一步分析。

8.7.2 示例:抗体浓度的测定

在两个实验室 A 和 B 中,用放射性免疫测定法测定 21 份血浆样品中特定抗体的浓度。每个实验室中的所有测量是对同一轮实验中的分割样品分别重复测量两次。所得浓度(以 U/L 为单位)如表 14 所示。由于被测量的取值范围很大,结果的差异也会很大,因此进行计算前对数据取自然对数。转化后的数据如表 15 所示,表中所给统计量的图示如图 12、图 13 和图 14。

由重复测定的极差图,可见实验室 A 重复测定的变异比实验室 B 大。这些统计量的联合值如表 15 所示,如有兴趣也可利用 F-检验进行比较。由图 14 可见,图中的点没有明显的模式或趋势。尽管图 12 和图 13 中重复测定的极差几乎都小于 0.2,但图 14 中大部分实验室间差值要比 0.2 大得多。这方面问题需要继续研究,因为这表明实验室间的差异更依赖于样品。实验室间平均之差的计算结果列在表 15 中,它表明实验室间差异的程度,但不可用于预测在之后的样品分析中的实验室间差异。因此,利用转化后的数据,平均来说, $\ln(B)-\ln(A)=0.443$ ,则  $B/A=1.6$ ,表明实验室 B 所得结果平均起来比实验室 A 多一个系数 1.6。然而,在某些样品中差值更大,而在其他样品中实验室 A 得到更大的结果。

表 14 21 份血浆样品中特定抗体的浓度

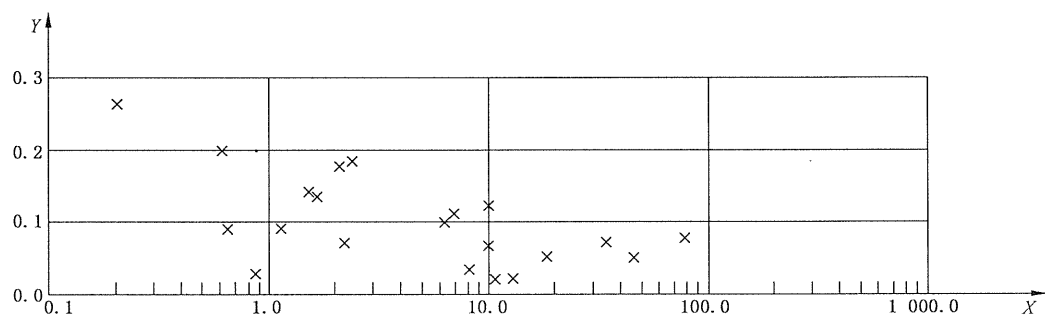
| 样本<br><i>i</i> | 实验室 A         |               | 实验室 B         |               | 实验室 A      | 实验室 B      | 实验室 A 和 B  |
|----------------|---------------|---------------|---------------|---------------|------------|------------|------------|
|                | 重复测量 1<br>U/L | 重复测量 2<br>U/L | 重复测量 1<br>U/L | 重复测量 2<br>U/L | 平均值<br>U/L | 平均值<br>U/L | 平均值<br>U/L |
| 1              | 19.106        | 18.174        | 11.473        | 11.705        | 18.640     | 11.589     | 15.115     |
| 2              | 6.424         | 7.171         | 5.812         | 5.812         | 6.798      | 5.812      | 6.305      |
| 3              | 6.619         | 5.989         | 11.705        | 11.473        | 6.304      | 11.589     | 8.947      |
| 4              | 0.543         | 0.664         | 0.861         | 0.905         | 0.604      | 0.883      | 0.743      |
| 5              | 43.816        | 46.063        | 49.899        | 55.147        | 44.940     | 55.523     | 48.731     |
| 6              | 2.096         | 2.535         | 24.047        | 26.843        | 2.316      | 25.445     | 13.880     |
| 7              | 10.591        | 9.875         | 9.116         | 8.671         | 10.233     | 8.894      | 9.563      |
| 8              | 13.874        | 13.599        | 12.554        | 12.807        | 13.737     | 12.681     | 13.209     |
| 9              | 1.974         | 2.363         | 1.094         | 1.020         | 2.169      | 1.057      | 1.613      |
| 10             | 9.393         | 10.591        | 13.736        | 14.585        | 9.992      | 14.161     | 12.076     |
| 11             | 1.840         | 1.616         | 2.484         | 2.460         | 1.728      | 2.472      | 2.100      |
| 12             | 31.817        | 34.124        | 48.424        | 55.147        | 32.971     | 51.786     | 42.378     |
| 13             | 1.150         | 1.051         | 2.014         | 2.270         | 1.101      | 2.142      | 1.621      |
| 14             | 0.625         | 0.684         | 1.051         | 1.174         | 0.655      | 1.113      | 0.884      |

表 14 (续)

| 样本<br><i>i</i>                                 | 实验室 A         |               | 实验室 B         |               | 实验室 A      | 实验室 B      | 实验室 A 和 B  |
|------------------------------------------------|---------------|---------------|---------------|---------------|------------|------------|------------|
|                                                | 重复测量 1<br>U/L | 重复测量 2<br>U/L | 重复测量 1<br>U/L | 重复测量 2<br>U/L | 平均值<br>U/L | 平均值<br>U/L | 平均值<br>U/L |
| 15                                             | 73.700        | 79.838        | 119.104       | 127.740       | 76.769     | 123.422    | 100.096    |
| 16                                             | 2.181         | 2.340         | 2.560         | 3.065         | 2.261      | 2.813      | 2.537      |
| 17                                             | 8.415         | 8.166         | 5.755         | 5.585         | 8.291      | 5.670      | 6.980      |
| 18                                             | 1.419         | 1.632         | 8.846         | 8.846         | 1.526      | 8.846      | 5.186      |
| 19                                             | 0.861         | 0.887         | 2.612         | 3.065         | 0.874      | 2.839      | 1.856      |
| 20                                             | 10.697        | 10.486        | 15.029        | 14.880        | 10.592     | 14.955     | 12.773     |
| 21                                             | 0.230         | 0.177         | 0.795         | 0.795         | 0.204      | 0.795      | 0.499      |
| 注：表中数据为每升(L)样品中千单位(kU)抗体数,其中单位(U)由国际标准物质的浓度定义。 |               |               |               |               |            |            |            |

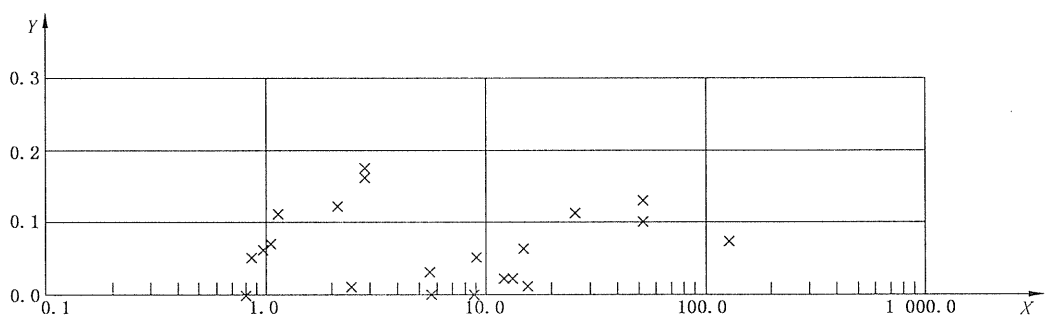
表 15 表 14 中数据的 ln(浓度)和统计量

| 样本<br><i>i</i>                                 | 实验室 A            |                  | 实验室 B            |                  | 实验室 A        | 实验室 B        | 差值 B-A<br>ln U/L |
|------------------------------------------------|------------------|------------------|------------------|------------------|--------------|--------------|------------------|
|                                                | 重复测量 1<br>ln U/L | 重复测量 2<br>ln U/L | 重复测量 1<br>ln U/L | 重复测量 2<br>ln U/L | 极差<br>ln U/L | 极差<br>ln U/L |                  |
| 1                                              | 2.95             | 2.90             | 2.44             | 2.46             | 0.05         | 0.02         | -0.475           |
| 2                                              | 1.86             | 1.97             | 1.76             | 1.76             | 0.11         | 0.00         | -0.155           |
| 3                                              | 1.89             | 1.79             | 2.46             | 2.44             | 0.10         | 0.02         | 0.610            |
| 4                                              | -0.61            | -0.41            | -0.15            | -0.10            | 0.20         | 0.05         | 0.385            |
| 5                                              | 3.78             | 3.83             | 3.91             | 4.01             | 0.05         | 0.10         | 0.155            |
| 6                                              | 0.74             | 0.93             | 3.18             | 3.29             | 0.19         | 0.11         | 2.400            |
| 7                                              | 2.36             | 2.29             | 2.21             | 2.16             | 0.07         | 0.05         | -0.140           |
| 8                                              | 2.63             | 2.61             | 2.53             | 2.55             | 0.02         | 0.02         | -0.080           |
| 9                                              | 0.68             | 0.86             | 0.09             | 0.02             | 0.18         | 0.07         | -0.715           |
| 10                                             | 2.24             | 2.36             | 2.62             | 2.68             | 0.12         | 0.06         | 0.350            |
| 11                                             | 0.61             | 0.48             | 0.91             | 0.90             | 0.13         | 0.01         | 0.360            |
| 12                                             | 3.46             | 3.53             | 3.88             | 4.01             | 0.07         | 0.13         | 0.450            |
| 13                                             | 0.14             | 0.05             | 0.70             | 0.82             | 0.09         | 0.12         | 0.665            |
| 14                                             | -0.47            | -0.38            | 0.05             | 0.16             | 0.09         | 0.11         | 0.530            |
| 15                                             | 4.30             | 4.38             | 4.78             | 4.85             | 0.08         | 0.07         | 0.475            |
| 16                                             | 0.78             | 0.85             | 0.94             | 1.12             | 0.07         | 0.18         | 0.215            |
| 17                                             | 2.13             | 2.10             | 1.75             | 1.72             | 0.03         | 0.03         | -0.380           |
| 18                                             | 0.35             | 0.49             | 2.18             | 2.18             | 0.14         | 0.00         | 1.760            |
| 19                                             | -0.15            | -0.12            | 0.96             | 1.12             | 0.03         | 0.16         | 1.175            |
| 20                                             | 2.37             | 2.35             | 2.71             | 2.70             | 0.02         | 0.01         | 0.345            |
| 21                                             | -1.47            | -1.73            | -0.23            | -0.23            | 0.26         | 0.00         | 1.371            |
| 联合值<br>实验室差值平均值                                |                  |                  |                  |                  | 0.119        | 0.083        | 0.443            |
| 注：表中数据为每升(L)样品中千单位(kU)抗体数,其中单位(U)由国际标准物质的浓度定义。 |                  |                  |                  |                  |              |              |                  |



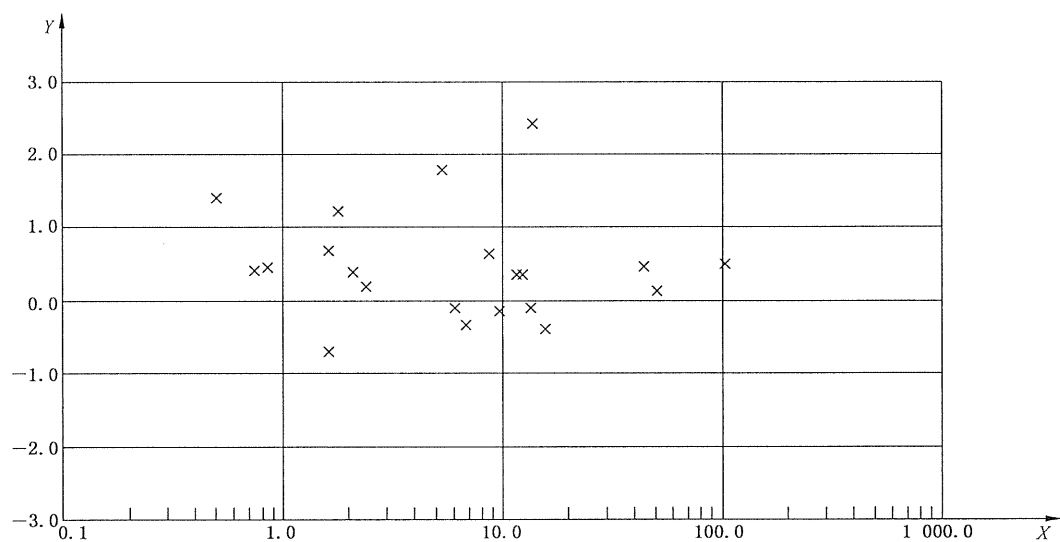
图例:  $X$  —— 实验室 A 的浓度平均值, %;  
 $Y$  —— 实验室 A 的  $\ln(\text{浓度})$  极差。

图 12 实验室 A 的重复测定极差(极差由  $\ln$  浓度计算)



图例:  $X$  —— 实验室 B 的浓度平均值, %;  
 $Y$  —— 实验室 B 的  $\ln(\text{浓度})$  极差。

图 13 实验室 B 的重复测定极差(极差由  $\ln$  浓度计算)



图例:  $X$  —— 实验室 A 和 B 的浓度平均值, %;  
 $Y$  —— 实验室 B 浓度的对数 - 实验室 A 浓度的对数。

图 14 实验室均值差  $B - A$  (浓度对数的差)

## 9 组合多轮能力验证计划中性能统计量值的图示法(见 GB/T 15483.1—1999, A.3.2)

### 9.1 应用

当要组合多轮计划的标准化性能统计量值时,协调者应考虑如 9.2 和 9.3 所述的图示法。这些图示组合了多轮计划的性能统计量值,可识别实验结果的趋势和其他特征,而这些特征在每轮计划得分的单独分析中是不明显的。

注:不推荐使用“运行分值”,即只计算实验室多轮计划获得的组合性能统计量值,而不在图中表示。由于实验室可能在某一轮计划中对某种物质的测试中出现“行动信号”或“警戒信号”,而在其他轮计划中没有,运行分值可能会掩盖这一情况。运行分值以发出行动和警戒信号次数表示(见 7.9)。

### 9.2 $z$ 值的常规控制图

#### 9.2.1 总则

要绘制  $z$  值的常规控制图,需将每个实验室所得  $z$  值单独绘出,且行动限和警戒限分别设在  $\pm 3.0$  和  $\pm 2.0$ ,具体形式如表 16 和图 15 所示。当一轮计划中需测量多个特性时,不同特性的  $z$  值可在同一图中绘出,但不同特征的点应使用不同的记号和/或颜色来表示。绘制常规控制图参见 GB/T 4091<sup>[4]</sup>。

常规控制图是一种能识别  $z$  值出现较大变异的有效方法。

常规控制图中,在下列情况下给出失控信号:

- 有一个点落在行动限( $\pm 3.0$ )外;
- 连续 3 个点中有两个落在相同的警戒限( $\pm 2.0$ )外。

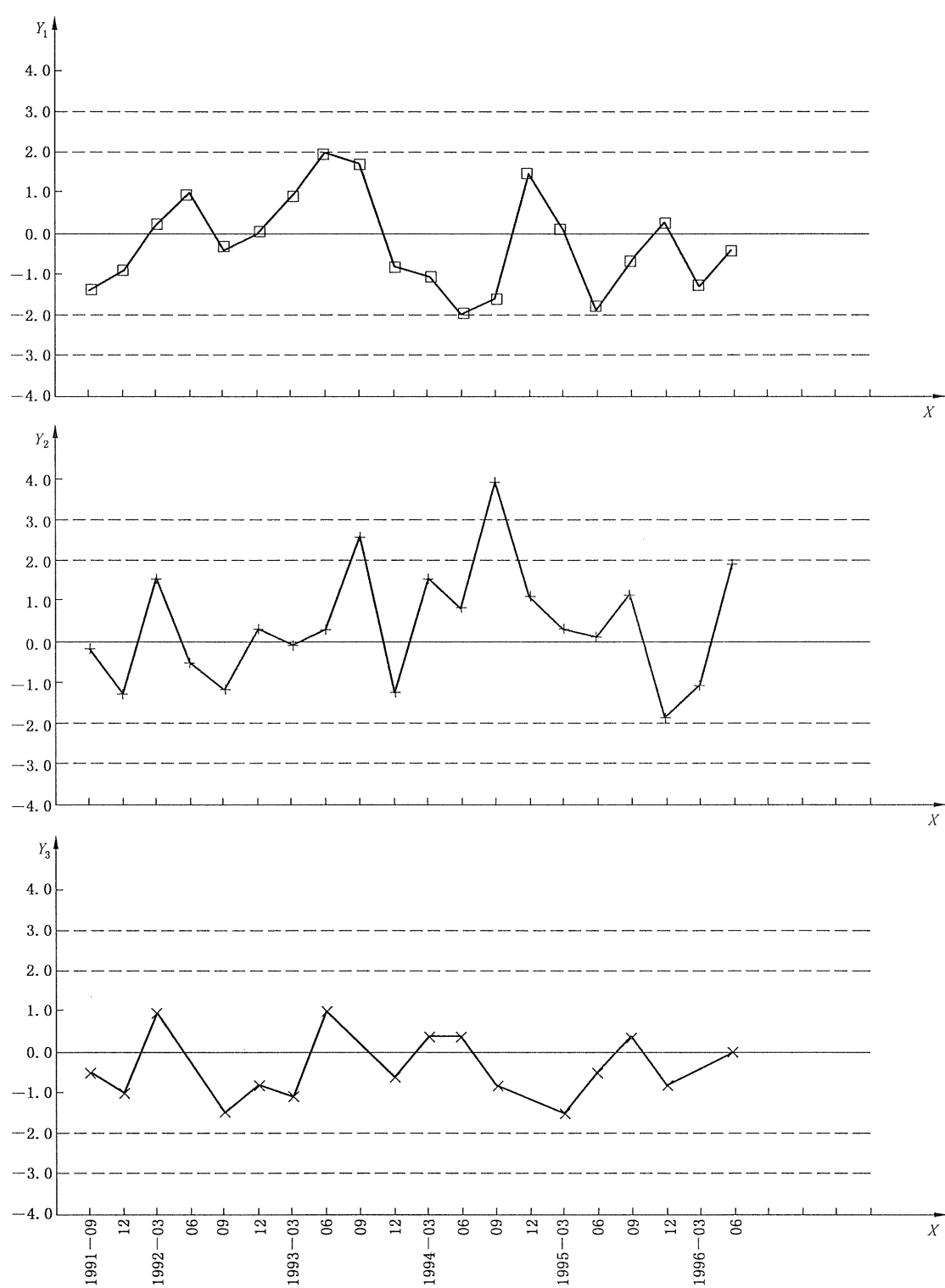
当常规控制图给出一个失控信号时,应按 4.1 中的有关条款采取行动。注意到能力评定标准差  $\hat{\sigma}$  不必是实验室偏倚的标准差,此时通常与常规控制图的行动限和警戒限相关的概率水平不适用。

#### 9.2.2 示例:过敏原浓度的测定

表 16 中数据在图 15 中绘出。

表 16 包含三种过敏原(d1,f1 和 e3)浓度的 20 轮能力验证计划所得数据

| 实验日期        | 过敏原的 $z$ 值 |      |      | $z$ 值的累积和 |      |      |
|-------------|------------|------|------|-----------|------|------|
|             | d1         | f1   | e3   | d1        | f1   | e3   |
| 1991 年 9 月  | -1.4       | -0.2 | -0.5 | -1.4      | -0.2 | -0.5 |
| 1991 年 12 月 | -0.9       | -1.3 | -1.0 | -2.3      | -1.5 | -1.5 |
| 1992 年 3 月  | 0.2        | 1.6  | 1.0  | -2.1      | 0.1  | -0.5 |
| 1992 年 6 月  | 1.0        | -0.5 |      | -1.1      | -0.4 |      |
| 1992 年 9 月  | -0.4       | -1.2 | -1.5 | -1.5      | -1.6 | -2.0 |
| 1992 年 12 月 | 0.0        | 0.3  | -0.8 | -1.5      | -1.3 | -2.8 |
| 1993 年 3 月  | 0.9        | -0.1 | -1.1 | -0.6      | -1.4 | -3.9 |
| 1993 年 6 月  | 2.0        | 0.3  | 1.0  | 1.4       | -1.1 | -2.9 |
| 1993 年 9 月  | 1.7        | 2.6  |      | 3.1       | 1.5  |      |
| 1993 年 12 月 | -0.8       | -1.3 | -0.6 | 2.3       | 0.2  | -3.5 |
| 1994 年 3 月  | -1.0       | 1.6  | 0.4  | 1.3       | 1.8  | -3.1 |
| 1994 年 6 月  | -2.0       | 0.8  | 0.4  | -0.7      | 2.6  | -2.7 |
| 1994 年 9 月  | -1.6       | 4.0  | -0.8 | -2.3      | 6.6  | -3.5 |
| 1994 年 12 月 | 1.5        | 1.1  |      | -0.8      | 7.7  |      |
| 1995 年 3 月  | 0.1        | 0.3  | -1.5 | -0.7      | 8.0  | -5.0 |
| 1995 年 6 月  | -1.9       | 0.1  | -0.5 | -2.6      | 8.1  | -5.5 |
| 1995 年 9 月  | -0.7       | 1.2  | 0.4  | -3.3      | 9.3  | -5.1 |
| 1995 年 12 月 | 0.3        | -1.9 | -0.8 | -3.0      | 7.4  | -5.9 |
| 1996 年 3 月  | -1.3       | -1.1 |      | -4.3      | 6.3  |      |
| 1996 年 6 月  | -0.4       | 1.9  | 0.0  | -4.7      | 8.2  | -5.9 |



图例: X —— 实验日期;  
 $Y_1$  —— 过敏原 d1 的 z 值;  
 $Y_2$  —— 过敏原 f1 的 z 值;  
 $Y_3$  —— 过敏原 e3 的 z 值。

图 15 一个实验室进行 20 轮能力验证计划所得数据的常规控制图(由表 16 数据)

9.3 z 值的累积和控制图

9.3.1 总则

要绘制  $z$  值的累积和控制图,需计算一个实验室在 20 轮能力验证计划中所得  $z$  值的累积和,并在图中描点,如图 16 所示。由于累积和图可用于检验  $z$  值是否偏离零点,故累积和的“目标值”为零,累积和可简单地由  $z$  值的累积和计算得到。绘制累积和图的方法见 GB/Z 4887—2006<sup>[5]</sup>。

当每轮计划中测量了多个特性时,应分别计算每个特性的累积和。

累积和控制图是一种在确定特性值时识别引起偏倚原因的有效方法,而这种偏倚持续存在于多轮计划中。当这样的累积和控制图给出了失控信号时,应按 4.1 中的有关条款采取行动。

9.3.2 示例:过敏原浓度的测定

表 16 包含了测定三种过敏原所得的  $z$  值及其累积和。累积和图如图 16 所示。由图形可知实验室结果从 1994 年 3 月开始发生变化。

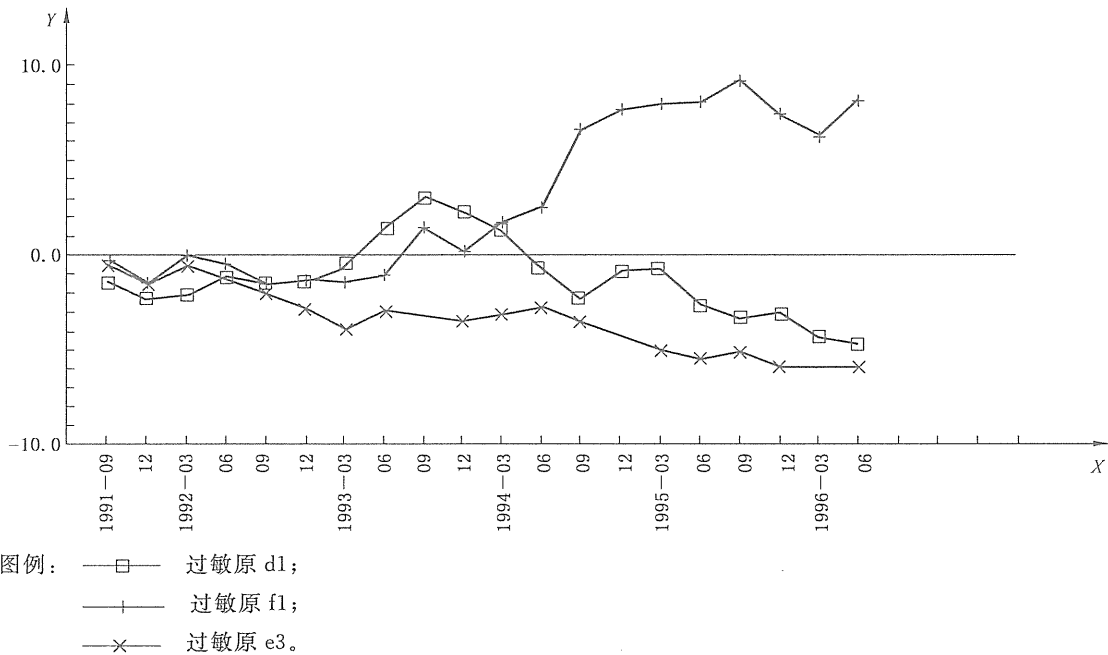


图 16 一个实验室在 20 轮能力验证计划中所得数据的累积和控制图(由表 16 数据)

9.4 实验室标准化偏倚对平均值的图

9.4.1 总则

当每轮能力验证计划的特性水平变化范围较大时, $z$  值或其他标准化性能统计量值关于特性水平的图将显示实验室偏倚是否随特性水平而变化。

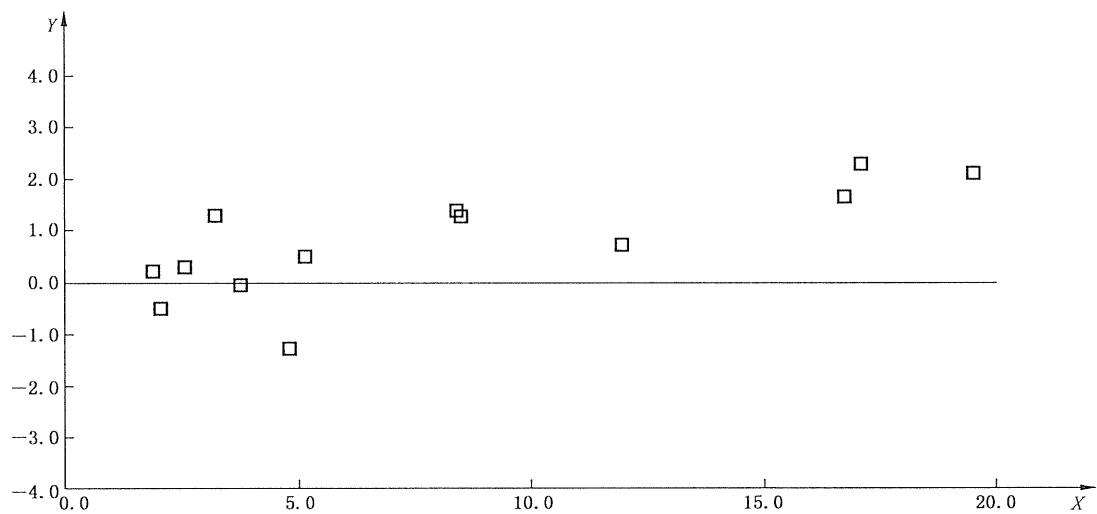
9.4.2 示例:过敏原浓度的测定

表 17 为一个实验室在 13 轮能力验证计划中报告的过敏原浓度数据,图 17 为其  $z$  值对浓度水平图。图中显示实验室在较高浓度水平有获得正的  $z$  值的趋势。对这些数据用回归分析可验证该趋势是统计显著的。

表 17 实验室测定的过敏原 d1 的浓度

| 日期          | 浓度平均值<br>kU/L | z 值  |
|-------------|---------------|------|
| 1992 年 6 月  | 12.00         | 0.7  |
| 1992 年 9 月  | 4.82          | -1.3 |
| 1992 年 12 月 | 17.20         | 2.3  |
| 1993 年 3 月  | 8.53          | 1.3  |
| 1993 年 6 月  | 8.47          | 1.4  |
| 1993 年 9 月  | 2.57          | 0.3  |
| 1993 年 12 月 | 1.94          | 0.2  |
| 1994 年 3 月  | 3.78          | -0.1 |
| 1994 年 6 月  | 2.07          | -0.5 |
| 1994 年 9 月  | 19.50         | 2.1  |
| 1994 年 12 月 | 3.13          | 1.3  |
| 1995 年 3 月  | 16.70         | 1.7  |
| 1995 年 6 月  | 5.12          | 0.5  |

注：表中数据为每升(L)样品中千单位(kU)抗体数,其中单位(U)由国际标准物质的浓度定义。



图例: X —— 浓度, kU/L;  
Y —— 过敏原 d1 的 z 值。

图 17 实验室测定过敏原 d1 的浓度水平与 z 值的关系

### 9.5 点图

当一轮能力验证计划中对同一个被测量有多个测试单元时,此时常规控制图等图方法仍然有用。点图与常规控制图相似。不同的是点图在每个时刻有多个点,并将每个时刻测量结果的平均值连线。

以下给出了每三个月对五份不同测试单元测定血糖浓度的示例。当指定值为 60 mg/dL 或更高

时,可接受的误差范围是指定值的 $\pm 10\%$ 内,而当指定值低于 60 mg/dL 时,可接受的误差范围是 $\pm 6.0$  mg/dL。这种情况以及指定值取值范围较大时,适合用  $z$  值进行评价,使用固定的能力评定标准差,根据指定值的大小,分别为 3.33%和 2.0 mg/dL。

图 18 显示在第一轮和第三轮计划中有个别结果落在 $\pm 3.0$  的边界之外,且序列中有一些可疑值。在这个数据集中,在第一轮计划中实验室重复性较差且偏倚较高。第二轮中精密度问题得到解决但偏倚未得到调整。第三轮显示已校正偏倚,但精密度问题再次出现。第四轮显示与前一轮有相似的偏倚,但精密度更好。

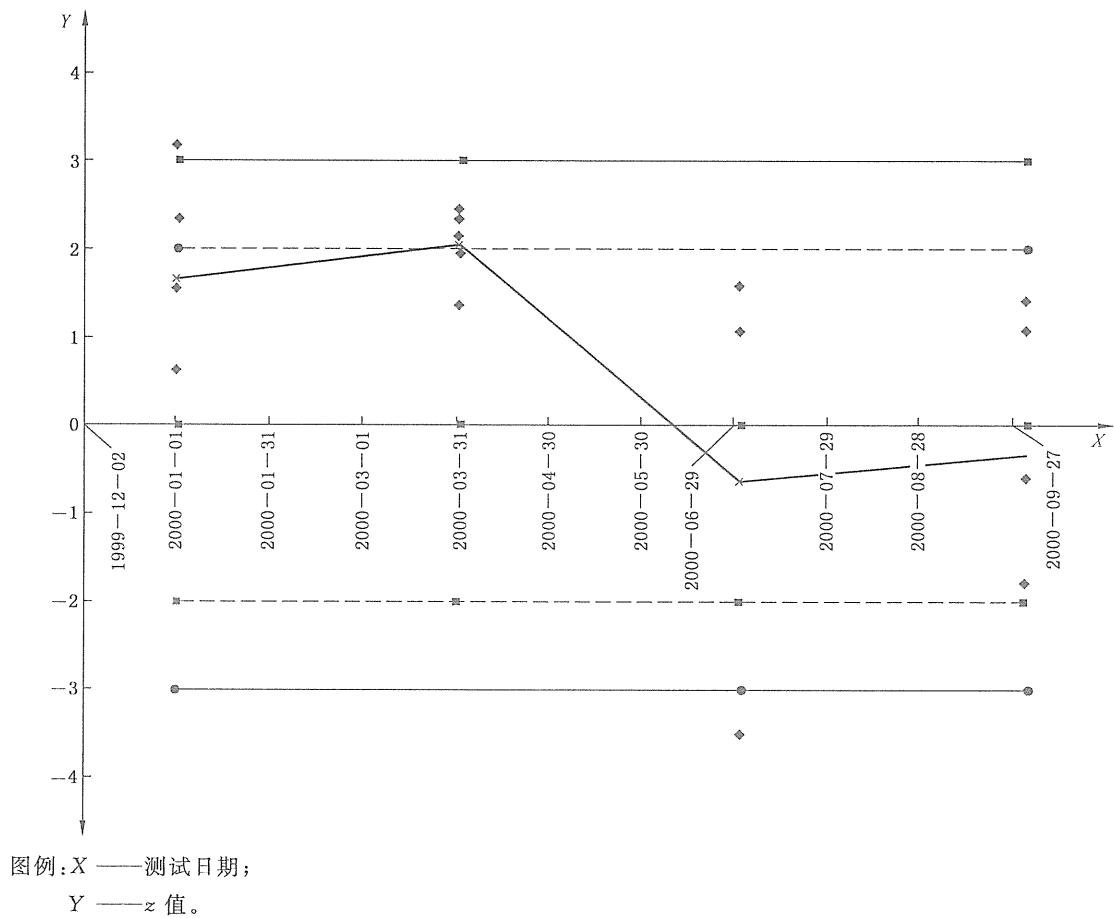


图 18 四轮能力验证计划中测定的血糖  $z$  值



附 录 A  
(规范性附录)  
符 号

- $g$  ——均匀性检验中测试的样本量；  
 $n$  ——每份样本的重复测量次数；  
 $p$  ——一轮能力验证计划中参与的实验室数；  
 $s_s$  ——样本间标准差；  
 $s_{\bar{x}}$  ——样本平均值的标准差；  
 $s^*$  ——稳健标准差(利用稳健算法计算的标准差)；  
 $u_X$  ——指定值的标准不确定度；  
 $w_t$  ——测试份样间极差；  
 $x$  ——测量结果；  
 $x^*$  ——稳健平均值(利用稳健算法计算的平均值)；  
 $z$  ——用于能力评估的性能统计量值；  
 $D$  ——实验室偏倚；  
 $X$  ——能力评估的指定值；  
 $\sigma_L$  ——实验室间标准差；  
 $\hat{\sigma}$  ——能力评定标准差；  
 $\sigma_r$  ——重复性标准差；  
 $\sigma_R$  ——再现性标准差；  
 $\phi$  ——验证由经验值得到  $\hat{\sigma}$  时所用的系数。

## 附 录 B

## (规范性附录)

## 样品的均匀性检验和稳定性检验

## B.1 均匀性检验程序

样品的均匀性检验程序如下：

- a) 选择一个实验室进行均匀性检验。当不需要对每个特性都进行均匀性检验时,可选择一种测量方法,以及一种对样品异质性比较敏感的特性进行均匀性检验。因此,对粒状物质,选择完全筛出特性(比如粒状物通过特定孔径筛网的百分比)。当测量结果为比值时,较小比值的特性的匀质化将更加困难,且在均匀性检验中会更容易检出不均匀性。(完全筛出特性是指单个粒子可显示特性的极端值。非完全筛出特性是单个粒子存在的可以变化的特性。)
- b) 制备并包装一轮能力验证计划的样品,应保证有充足的样品供给参加者以及均匀性检验。
- c) 在最终包装的样品中随机选择  $g$  份,  $g \geq 10$ 。若可从前期对相似样品利用相同程序进行的均匀性检验中得到适合的数据,可减少均匀性检验的样本量。
- d) 从一个样品中制备两个测试份量,选择适合的方法将测试份量间差异降到最小。
- e) 将  $2g$  份测试份量随机排序,得到每份的测量结果,整个测量在重复性条件下完成。
- f) 计算总平均值  $\bar{x}_{..}$ , 样品内标准差  $s_w$  和样品间标准差  $s_s$ , 如 B.3 所示。

## B.2 均匀性检验的评估标准

比较样品间标准差  $s_s$  和能力评定标准差  $\hat{\sigma}$ 。当

$$s_s \leq 0.3\hat{\sigma} \quad \dots\dots\dots (B.1)$$

可认为样品是充分均匀的。式(B.1)中系数取为 0.3 是因为,当以上条件满足时,样品间方差对能力评定方差的贡献不多于 10%。若以上条件不满足,协调者应考虑以下可能情况:

- a) 检查样品制备程序是否可以改进。
- b) 将一定数量的样品分发给能力验证计划的参加者,并要求得到每份样品的测试结果。样品的异质性将使样品内标准差增加到:

$$\sigma_{r1} = \sqrt{\sigma_r^2 + s_s^2} \quad \dots\dots\dots (B.2)$$

此时,利用式(2)确定重复测试次数时,用  $\sigma_{r1}$  代替  $\sigma_r$ 。

- c) 在能力评定标准差中包含了样品间标准差,利用计算  $\hat{\sigma}$  的公式

$$\hat{\sigma} = \sqrt{\hat{\sigma}_1^2 + s_s^2} \quad \dots\dots\dots (B.3)$$

其中  $\hat{\sigma}_1$  是样本中不含异质成分时的能力评定标准差。

## B.3 均匀性检验的公式

均匀性检验的数据表示为:

$$x_{t,k}$$

其中

$t$  表示样品 ( $t=1, 2, \dots, g$ )

$k$  表示测试份量( $k=1,2$ )。

定义样品均值为:

$$\bar{x}_{t..} = (x_{t,1} + x_{t,2})/2 \quad \dots\dots\dots (B.4)$$

测试份量间极差为:

$$w_t = |x_{t,1} - x_{t,2}| \quad \dots\dots\dots (B.5)$$

计算总平均值为:

$$\bar{x}_{...} = \sum \bar{x}_{t..} / g \quad \dots\dots\dots (B.6)$$

样品均值的标准差为:

$$s_{\bar{x}} = \sqrt{\sum (\bar{x}_{t..} - \bar{x}_{...})^2 / (g-1)} \quad \dots\dots\dots (B.7)$$

样品内标准差为:

$$s_w = \sqrt{\sum w_t^2 / (2g)} \quad \dots\dots\dots (B.8)$$

其中对所有样品求和( $t=1,2,\dots,g$ )。

最后,计算样品间标准差为:

$$s_s = \sqrt{s_{\bar{x}}^2 - (s_w^2/2)} \quad \dots\dots\dots (B.9)$$

注:可用测试份量间标准差代替极差,如  $s_t = w_t / \sqrt{2}$ 。

#### B.4 稳定性检验的程序

- a) 使用与均匀性检验相同的实验室进行稳定性检验。使用相同的测量方法,并测量相同的材料特性。
- b) 允许均匀性检验和稳定性检验间有时间延迟,这与能力验证中参加者检测样本的时间延迟是相似的。
- c) 随机选取  $g$  个样品,其中  $g \geq 3$ 。
- d) 使用与均匀性检验相同的方法,从每一份样品中制备两个测试份量。
- e) 将  $2g$  个测试份量按随机顺序测量,所有测量在重复性条件下完成,每个测试份量得到一个测量结果  $y_{t,k}$ 。
- f) 计算稳定性检验中测量结果的总平均值  $y_{...}$ 。

#### B.5 稳定性检验的评估准则

比较均匀性检验中测量结果的总平均值与稳定性检验中测量结果的总平均值,如果

$$|\bar{x}_{...} - \bar{y}_{...}| \leq 0.3\hat{\sigma} \quad \dots\dots\dots (B.10)$$

可认为样品是充分稳定的,如果以上条件不满足,则检查样品制备和储存程序是否可以改进。

#### B.6 示例:大豆粉中的铜含量( $\mu\text{g/g}$ )

在能力验证计划中检测大豆粉中的铜含量,根据精密度试验结果,将能力评定标准差设定为  $\hat{\sigma} = 1.1 \mu\text{g/g}$ 。根据准则(B.1),样品间标准差不应大于

$$0.3 \times 1.1 = 0.330 \mu\text{g/g}$$

对于均匀性检验,随机选取 12 份能力验证计划中准备的大豆粉样品,且每份样品分成两个测试份量,确定其中的铜含量。表 B.1 给出了这些数据,以及样品平均值和份量间极差。

根据 B.3 中公式可计算以下值:

总平均值  $\bar{x}_{...} = 10.02 \mu\text{g/g}$

样品均值标准差  $s_{\bar{x}} = 0.340 \mu\text{g/g}$

样品内标准差  $s_w = 0.246 \mu\text{g/g}$

因此，

样品间标准差  $s_s = 0.292 \mu\text{g/g}$

由于这个值小于  $0.330 \mu\text{g/g}$ ，可推断能力验证计划中所用样品是充分均匀的。

对于稳定性检验，其中三份样本已在一个月后在同一个实验室中检测，其结果的总平均值为  $\bar{y}_{...} = 10.78 \mu\text{g/g}$ 。差值  $\bar{y}_{...} - \bar{x}_{...} = 0.76 \mu\text{g/g}$  超过了  $0.3\hat{\sigma} = 0.33 \mu\text{g/g}$ ，因此可以推断出样本不是充分稳定的。

表 B.1 均匀性检验的测试结果，大豆粉中的铜含量(mg/g)

| 样本编号<br>$t$ | 测试份量<br>1 | 测试份量<br>2 | 样本平均值<br>$x_{t_r}$ | 测试单元间极差<br>$w_t$ |
|-------------|-----------|-----------|--------------------|------------------|
| 1           | 10.5      | 10.4      | 10.45              | 0.1              |
| 2           | 9.6       | 9.5       | 9.55               | 0.1              |
| 3           | 10.4      | 9.9       | 10.15              | 0.5              |
| 4           | 9.5       | 9.9       | 9.70               | 0.4              |
| 5           | 10.0      | 9.7       | 9.85               | 0.3              |
| 6           | 9.6       | 10.1      | 9.85               | 0.5              |
| 7           | 9.8       | 10.4      | 10.10              | 0.6              |
| 8           | 9.8       | 10.2      | 10.00              | 0.4              |
| 9           | 10.8      | 10.7      | 10.75              | 0.1              |
| 10          | 10.2      | 10.0      | 10.10              | 0.2              |
| 11          | 9.8       | 9.5       | 9.65               | 0.3              |
| 12          | 10.2      | 10.0      | 10.10              | 0.2              |

## 附 录 C

### (规范性附录)

### 稳 健 分 析

#### C.1 算法 A

应用本算法可以得到数据平均值和标准差的稳健值。

注 1: 本附录重复了 GB/T 6379.5 中给出的算法 A 和算法 S。

注 2: 稳健性是估计算法的特点,而不是其产生的估计值的特点,因此严格来说,称由此算法计算的平均值和标准差是稳健的是不确切的。然而,为避免使用繁琐的术语,本标准中的“稳健平均值”和“稳健标准差”应理解为利用稳健算法计算的总体平均值和总体标准差的均值估计。

按递增顺序排列  $p$  个数据,表示为:

$x_1, x_2, \dots, x_i, \dots, x_p$ 。

这些数据的稳健平均值和稳健标准差记为  $x^*$  和  $s^*$ 。

计算  $x^*$  和  $s^*$  的初始值如下(med 表示中位数):

$$x^* = \text{med} x_i \quad (i=1, 2, \dots, p) \quad \dots\dots\dots (C.1)$$

$$s^* = 1.483 \times \text{med} |x_i - x^*| \quad (i=1, 2, \dots, p) \quad \dots\dots\dots (C.2)$$

根据以下步骤更新  $x^*$  和  $s^*$  的值。计算:

$$\delta = 1.5s^* \quad \dots\dots\dots (C.3)$$

对每个  $x_i (i=1, 2, \dots, p)$ , 计算

$$x_i^* = \begin{cases} x^* - \delta, & \text{若 } x_i < x^* - \delta \\ x^* + \delta, & \text{若 } x_i > x^* + \delta \\ x_i, & \text{其他} \end{cases} \quad \dots\dots\dots (C.4)$$

再由下式计算  $x^*$  和  $s^*$  的新的取值:

$$x^* = \sum x_i^* / p \quad \dots\dots\dots (C.5)$$

$$s^* = 1.134 \sqrt{\sum (x_i^* - x^*)^2 / (p-1)} \quad \dots\dots\dots (C.6)$$

其中求和符号对  $i$  求和。

稳健估计值  $x^*$  和  $s^*$  可由迭代计算得出,例如用已修改数据更新  $x^*$  和  $s^*$ ,直至过程收敛。当稳健标准差的第三位有效数字和稳健平均值相对应的数字在连续两次迭代中不再变化时,即可认为过程是收敛的。这是一种可用计算机编程实现的简单方法。

#### C.2 算法 S

此算法用于标准差(或极差),可推出标准差或极差的稳健联合值。

将  $p$  个数据以递增顺序排列,表示为:

$w_1, w_2, \dots, w_i, \dots, w_p$ 。

(这些数据可以是极差或标准差。)

稳健联合值记为  $w^*$ ,每个  $w_i$  相关的自由度为  $v$ 。(当  $w_i$  为极差时,  $v=1$ 。当  $w_i$  为  $n$  次测试结果的标准差时,  $v=n-1$ 。)根据表 C.1,查得算法所需的  $\xi$  和  $\eta$  值。

计算  $w^*$  的初始值如下 (med 表示中位数),  
 $w^* = \text{med} w_i \quad (i = 1, 2, \dots, p) \dots\dots\dots (C.7)$

按以下步骤更新  $w^*$  的值, 计算  
 $\phi = \eta \times w^* \dots\dots\dots (C.8)$

对于每个  $w_i (i = 1, 2, \dots, p)$ , 计算  
 $w_i^* = \begin{cases} \phi, & \text{若 } w_i > \phi \\ w_i, & \text{其他} \end{cases} \dots\dots\dots (C.9)$

计算新的  $w^*$  :  
 $w^* = \xi \sqrt{\sum (w_i^*)^2 / p} \dots\dots\dots (C.10)$

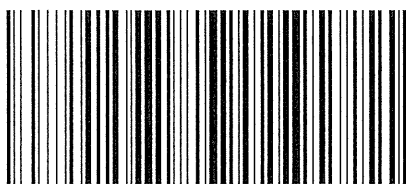
稳健估计值  $w^*$  可由迭代算法得到, 即不断更新  $w^*$ , 直到过程收敛。当稳健估计值的第三位有效数字连续两次迭代后数值不再变化时, 即可认为过程是收敛的。这是一种可利用计算机编程实现的简单方法。

表 C.1 稳健分析必需的因子: 算法 S

| 自由度<br>$v$                                       | 限系数<br>$\eta$ | 修正系数<br>$\xi$ |
|--------------------------------------------------|---------------|---------------|
| 1                                                | 1.645         | 1.097         |
| 2                                                | 1.517         | 1.054         |
| 3                                                | 1.444         | 1.039         |
| 4                                                | 1.395         | 1.032         |
| 5                                                | 1.359         | 1.027         |
| 6                                                | 1.332         | 1.024         |
| 7                                                | 1.310         | 1.021         |
| 8                                                | 1.292         | 1.019         |
| 9                                                | 1.277         | 1.018         |
| 10                                               | 1.264         | 1.017         |
| 注: $\eta$ 和 $\xi$ 的值由 GB/T 6379.5—2006 的附录 B 导出。 |               |               |

## 参 考 文 献

- [1] Thompson, M and Wood, R. The international harmonized protocol for the proficiency testing of (chemical) analytical laboratories, AOAC, September 1993
- [2] Jackson, J. E. Quality control methods for two related variables, Industrial Quality Control, January, 1956
- [3] Horwitz, W. Evaluation of analytical methods used for regulations of food and drugs. Anal. Chem., 54, 1982, pp. 67A-76A
- [4] GB/T 4091—2001 常规控制图
- [5] GB/Z 4887—2006 累积和图——运用累积和技术进行质量控制和数据分析指南
- [6] Van Nuland, Y. ISO 9002 and the circle technique, Qual. Eng., 5, 1992, pp. 269-291
- [7] GB/T 6379.1—2004 测量方法与结果的准确度(正确度与精密度) 第1部分:总则与定义
- [8] GB/T 6379.2—2004 测量方法与结果的准确度(正确度与精密度) 第2部分:确定标准测量方法重复性与再现性的基本方法
- [9] ISO 5725-3 测量方法与结果的准确度(正确度与精确度) 第3部分:标准测量方法精密度的中间度量
- [10] GB/T 6379.4—2006 测量方法与结果的准确度(正确度与精确度) 第4部分:确定标准测量方法正确度的基本方法
- [11] GB/T 6379.5—2006 测量方法与结果的准确度(正确度与精确度) 第5部分:确定标准测量方法正确度的可替代方法
- [12] GB/T 6379.6—2009 测量方法与结果的准确度(正确度与精确度) 第6部分:准确度值的实际应用
- [13] GB/T 27025—2008 检测和校准实验室能力的通用要求



GB/T 28043-2011

版权专有 侵权必究

\*

书号:155066·1-44297